

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 16, 2003, 10:38:31 ; Search time 29 Seconds

(without alignments)
85.261 Million cell updates/sec

Title: US-09-551-151a-43

Perfect score: 64

Sequence: 1 SPQGIAGRNFN 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 500 summaries

Database :
1: SP_ARCHA:21:
2: SP_BACTERIA:
3: SP_FUNGI:
4: SP_HUMAN:
5: SP_INVERTEBRATE:
6: SP_MAMMAL:
7: SP_MHC:
8: SP_ORGANELLE:
9: SP_PHAGE:
10: SP_PLANT:
11: SP_PROTOZOA:
12: SP_VIRUS:
13: SP_VIRUS:
14: SP_UNCLASSIFIED:
15: SP_VIRUS:
16: SP_BACTERIA:
17: SP_ARCHA:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	65.6	230	11 Q9R149	G9R149 cavia porce
2	42	65.6	589	11 Q991L6	G991L6 mus musculu
3	42	65.6	1450	13 Q9Y1B4	G9Y1B4 cynops pyrr
4	42	65.6	1453	11 Q63079	G63079 rattus norv
5	42	65.6	1461	4 Q76045	Q76045 homo sapien
6	40	62.5	113	11 Q8V172	Q8V172 cavia porce
7	40	62.5	347	6 Q9X125	Q9X125 bos taurus
8	40	62.5	475	10 Q80592	Q80592 arabisopsis
9	40	62.5	1160	4 Q14046	Q14046 homo sapien
10	40	62.5	1418	13 Q9W7R9	Q9W7R9 cynops pyrr
11	40	62.5	1419	11 Q63123	Q63123 rattus norv
12	40	62.5	1420	13 Q90W37	Q90W37 gallus gall
13	40	62.5	1442	11 Q62031	Q62031 mus musculu
14	40	62.5	1442	11 Q62033	Q62033 mus musculu
15	40	62.5	1459	11 Q62032	Q62032 mus musculu
16	40	62.5	1459	11 Q62032	Q62032 mus musculu

17	40	62.5	1487	4 Q14047	Q14047 homo sapien
18	40	62.5	1487	6 Q77753	Q77753 canis famli
19	39	60.9	211	2 Q9K17	Q9K17 ristaonia s
20	38	59.4	373	2 Q9K17	Q9K17 yersinia en
21	38	59.4	373	2 Q9K17	Q9K17 yersinia en
22	38	59.4	695	10 Q82804	Q82804 arabisopsis
23	38	59.4	702	10 Q04419	Q04419 arabisopsis
24	38	59.4	809	13 Q93485	Q93485 oncorhynch
25	38	59.4	1449	13 Q910C0	Q910C0 oncorhynch
26	38	59.4	1647	16 Q9A180	Q9A180 streptococ
27	37.5	58.6	441	5 Q9U9P3	Q9U9P3 streptococ
28	37.5	58.6	953	5 Q9Y125	Q9Y125 streptococ
29	37	57.8	84	16 Q97550	Q97550 streptococ
30	37	57.8	328	4 Q95061	Q95061 homo sapien
31	37	57.8	463	5 Q8T0H4	Q8T0H4 notophthal
32	37	57.8	496	13 Q42434	Q42434 notophthal
33	37	57.8	739	16 Q8XDP3	Q8XDP3 escherich
34	37	57.8	771	16 Q8U724	Q8U724 agrobacteri
35	37	57.8	865	5 Q9V6K7	Q9V6K7 agrobacteri
36	37	57.8	1158	2 Q93EC2	Q93EC2 rhizobium
37	37	57.8	1486	13 Q91717	Q91717 xenopus lae
38	36.5	57.0	365	5 Q5769	Q5769 caenorhabdi
39	36	56.2	100	4 Q9Y3P3	Q9Y3P3 homo sapien
40	36	56.2	149	16 Q915P4	Q915P4 pseudomonas
41	36	56.2	409	11 Q920V4	Q920V4 mus musculu
42	36	56.2	409	11 Q920V3	Q920V3 mus musculu
43	36	56.2	409	11 Q920V2	Q920V2 mus musculu
44	36	56.2	409	11 Q91VE9	Q91VE9 mus spicill
45	36	56.2	409	11 Q91V58	Q91V58 mus musculu
46	36	56.2	409	11 Q91V19	Q91V19 mus musculu
47	36	56.2	422	11 Q54905	Q54905 mus musculu
48	36	56.2	436	3 Q9P8N1	Q9P8N1 coriolus ve
49	36	56.2	441	4 Q96A84	Q96A84 homo sapien
50	36	56.2	491	2 Q9KWS5	Q9KWS5 pseudomonas
51	36	56.2	531	4 Q9BZG5	Q9BZG5 homo sapien
52	36	56.2	532	17 Q9YU22	Q9YU22 pyrococcus
53	36	56.2	539	4 Q9NUA2	Q9NUA2 homo sapien
54	36	56.2	542	4 Q9BZG6	Q9BZG6 homo sapien
55	36	56.2	544	4 Q9BZG7	Q9BZG7 homo sapien
56	36	56.2	556	4 Q9Y6M1	Q9Y6M1 homo sapien
57	36	56.2	598	16 Q8U6L9	Q8U6L9 agrobacteri
58	36	56.2	643	4 Q8TDS1	Q8TDS1 homo sapien
59	36	56.2	657	16 Q9ZLH1	Q9ZLH1 helicobacte
60	36	56.2	660	16 Q25414	Q25414 helicobacte
61	36	56.2	668	16 Q9ZKR5	Q9ZKR5 helicobacte
62	36	56.2	730	4 Q13771	Q13771 homo sapien
63	36	56.2	777	16 Q8ZLA2	Q8ZLA2 salmonella
64	36	56.2	777	16 Q8ZLA2	Q8ZLA2 salmonella
65	36	56.2	826	10 Q41059	Q41059 pisum sativ
66	36	56.2	830	10 Q04864	Q04864 solanum tub
67	36	56.2	832	10 Q9M6P8	Q9M6P8 sorghum bic
68	36	56.2	847	10 Q9XISA	Q9XISA phaeocystis
69	36	56.2	852	10 Q08131	Q08131 manihot esc
70	36	56.2	906	4 Q9UN21	Q9UN21 homo sapien
71	36	56.2	1445	13 Q9J251	Q9J251 rana catesb
72	36	56.2	1447	13 Q9IB91	Q9IB91 xenopus lae
73	36	56.2	1779	5 Q9VMV4	Q9VMV4 dirosophila
74	35	54.7	23	2 Q9KWN9	Q9KWN9 streptococ
75	35	54.7	89	2 Q54470	Q54470 staphylococ
76	35	54.7	182	5 Q9XWC2	Q9XWC2 caenorhabdi
77	35	54.7	259	11 Q99U68	Q99U68 mus musculu
78	35	54.7	369	5 Q21863	Q21863 caenorhabdi
79	35	54.7	393	11 Q9J9H7	Q9J9H7 mus musculu
80	35	54.7	396	11 Q9R084	Q9R084 rattus norv
81	35	54.7	396	11 Q9OXL3	Q9OXL3 rattus norv
82	35	54.7	422	4 Q9BXC9	Q9BXC9 homo sapien
83	35	54.7	422	4 Q43825	Q43825 homo sapien
84	35	54.7	437	16 Q97MC0	Q97MC0 clostridium
85	35	54.7	548	12 Q83880	Q83880 norway vir
86	35	54.7	548	12 Q68291	Q68291 human calli
87	35	54.7	562	12 Q90RA3	Q90RA3 tomato ring
88	35	54.7	562	12 Q90RA3	Q90RA3 tomato ring
89	35	54.7	564	12 Q9TFA2	Q9TFA2 tomato ring

90	35	54.7	605	4	Q96MB9	Q96mb9 homo sapien	163	34	53.1	947	16	Q8UEN6	Q8uen6 agrobacteri
91	35	54.7	618	11	Q9R083	Q9r083 rattus norv	164	34	53.1	947	16	Q92037	Q92q37 rhizobium m
92	35	54.7	618	11	Q9QXL4	Q9qxl4 rattus norv	165	34	53.1	947	16	Q92037	Q92q37 caenorhabdi
93	35	54.7	619	11	Q8VDJ4	Q8vdj4 mus musculu	166	34	53.1	1051	5	Q26055	Q26055 paracentrot
94	35	54.7	633	16	Q9KYJ7	Q9kyj7 streptomyc	167	34	53.1	1060	10	Q9LFT2	Q9lft2 arabadopsi
95	35	54.7	656	12	Q9DMC3	Q9dmc3 rat ctyomeg	168	34	53.1	1158	16	Q8Y2U0	Q8yx20 anabaena sp
96	35	54.7	675	5	Q9NS13	Q9ns13 caenorhabdi	169	34	53.1	1298	14	Q8WX42	Q8wx42 homo sapien
97	35	54.7	678	13	Q93486	Q93486 concorhynch	170	34	53.1	2459	16	Q8YTR8	Q8ytr8 anabaena sp
98	35	54.7	711	13	Q91BC7	Q91bc7 xenopus lae	171	33	52.3	298	10	Q9FNN6	Q9fnn6 arabadopsi
99	35	54.7	716	13	Q9PDM1	Q9pdm1 xenopus lae	172	33	52.3	826	10	Q940X9	Q940x9 arabadopsi
100	35	54.7	721	16	Q9KDS7	Q9kds7 bacillus ha	173	33	51.6	145	11	Q64Z98	Q64z98 rattus norv
101	35	54.7	807	10	Q9XED2	Q9xed2 triticum ae	174	33	51.6	147	4	Q96N53	Q96n53 homo sapien
102	35	54.7	810	10	Q9XGB3	Q9xgb3 triticum ae	175	33	51.6	149	2	Q93P56	Q93p56 comamonas t
103	35	54.7	823	10	Q41740	Q41740 zea mays (m	176	33	51.6	149	16	Q8RNY0	Q8rny0 comamonas t
104	35	54.7	829	10	Q9XED3	Q9xed3 agelipos ta	177	33	51.6	149	16	Q92068	Q92068 rhizobium m
105	35	54.7	830	10	Q04074	Q04074 triticum ae	178	33	51.6	158	16	Q98F46	Q98f46 rhizobium l
106	35	54.7	833	10	Q9XGB1	Q9xgb1 triticum ae	179	33	51.6	167	17	Q8RYP6	Q8ryp6 methanopyru
107	35	54.7	833	10	Q9FUB8	Q9fub8 triticum ae	180	33	51.6	172	16	Q8RAD5	Q8rad5 thermomane
108	35	54.7	865	10	Q9XGB2	Q9xgb2 triticum ae	181	33	51.6	192	5	Q9VVA3	Q9vva3 drosophila
109	35	54.7	895	6	Q9GKN9	Q9gkn9 sus scrofa	182	33	51.6	207	3	Q96U05	Q96u05 neuropep
110	35	54.7	896	6	Q9GKL7	Q9gkl7 sus scrofa	183	33	51.6	228	2	Q9KJN5	Q9kjin5 myxococcu
111	35	54.7	914	4	Q9UPW0	Q9upw0 homo sapien	184	33	51.6	239	10	Q23959	Q23959 glycine max
112	35	54.7	924	16	Q915U2	Q915u2 pseudomonas	185	33	51.6	255	11	Q08745	Q08745 mus musculu
113	35	54.7	924	16	Q915U2	Q915u2 pseudomonas	186	33	51.6	256	11	Q62361	Q62361 mus musculu
114	35	54.7	1058	13	Q9YGF0	Q9ygf0 oncorhynch	187	33	51.6	260	9	Q8SD18	Q8sd18 pseudomonas
115	35	54.7	1059	13	Q9YGF0	Q9ygf0 oncorhynch	188	33	51.6	261	16	Q54172	Q54172 streptomyc
116	35	54.7	1194	4	Q9H7K4	Q9h7k4 homo sapien	189	33	51.6	284	10	P93506	P93506 triticiu com
117	35	54.7	1458	13	Q910B9	Q910b9 oncorhynch	190	33	51.6	291	16	Q8ZMP4	Q8zmp4 salmonella
118	35	54.7	3198	5	Q26639	Q26639 strongyloce	191	33	51.6	293	10	Q9XEX4	Q9xex4 arabadopsi
119	35	54.7	4283	13	Q8UML7	Q8uml7 figu rubrip	192	33	51.6	297	11	Q08743	Q08743 mus musculu
120	34	53.1	340	10	Q9PLP7	Q9plp7 arabadopsi	193	33	51.6	304	11	Q922Y6	Q922y6 mus musculu
121	34	53.1	49	16	Q9K2C6	Q9k2c6 chlamydia p	194	33	51.6	304	11	Q08744	Q08744 mus musculu
122	34	53.1	119	17	Q9Y9V6	Q9y9v6 aeopyrium p	195	33	51.6	310	13	Q90612	Q90612 gallus gall
123	34	53.1	125	16	Q8DWM3	Q8dwm3 yerishia pe	196	33	51.6	321	16	Q8YEM2	Q8yem2 anabaena sp
124	34	53.1	154	2	Q22591	Q22591 caenorhabdi	197	33	51.6	322	11	Q922Y5	Q922y5 mus musculu
125	34	53.1	240	4	Q8WME2	Q8wme2 homo sapien	198	33	51.6	350	11	Q923Y6	Q923y6 mus musculu
126	34	53.1	263	16	Q9KA41	Q9ka41 bacillus ha	199	33	51.6	362	16	Q9PC05	Q9pc05 xylella fas
127	34	53.1	264	2	Q52007	Q52007 vibrio fisc	200	33	51.6	370	10	Q9VY98	Q9vy98 polytomella
128	34	53.1	264	16	Q92000	Q92q00 rhizobium m	201	33	51.6	381	16	Q92UX4	Q92ux4 rhizobium m
129	34	53.1	279	16	Q8Y0Z4	Q8ygz4 anabaena sp	202	33	51.6	388	2	P93561	P93561 solanum tub
130	34	53.1	280	16	Q8YWS3	Q8yws3 anabaena sp	203	33	51.6	397	2	Q8RMP1	Q8rmp1 pseudomonas
131	34	53.1	286	16	Q9PR14	Q9pr14 ureaplasma	204	33	51.6	397	2	Q8RL91	Q8rl91 pseudomonas
132	34	53.1	290	5	Q26054	Q26054 paracentrot	205	33	51.6	397	2	Q8RL91	Q8rl91 pseudomonas
133	34	53.1	370	10	Q43458	Q43458 glycine max	206	33	51.6	403	12	Q9YML3	Q9ylm3 lymantria d
134	34	53.1	393	16	Q9HV27	Q9hv27 pseudomonas	207	33	51.6	442	10	Q8S1U6	Q8slu6 oryza sativ
135	34	53.1	395	3	Q8XOK9	Q8xok9 neuropep	208	33	51.6	445	5	Q18812	Q18812 caenorhabdi
136	34	53.1	420	16	Q926G0	Q926g0 rhizobium m	209	33	51.6	461	16	Q9PGF5	Q9pgf5 xylella fas
137	34	53.1	425	16	Q9KYE9	Q9kye9 streptomyc	210	33	51.6	465	16	Q8YCH2	Q8ych2 bruceella me
138	34	53.1	429	16	Q91006	Q91006 pseudomonas	211	33	51.6	470	16	Q8ZM23	Q8zm23 salmonella
139	34	53.1	481	2	Q9S345	Q9s345 prevotella	212	33	51.6	470	16	Q8Z3R7	Q8z3r7 salmonella
140	34	53.1	481	10	Q9LNT1	Q9lnl1 arabadopsi	213	33	51.6	481	10	P92934	P92934 arabadopsi
141	34	53.1	503	16	Q8LRH0	Q8lrh0 salmonella	214	33	51.6	481	16	Q8UBD7	Q8ubd7 agrobacteri
142	34	53.1	503	16	Q8LRH0	Q8lrh0 salmonella	215	33	51.6	483	13	Q13045	Q13045 xenopus lae
143	34	53.1	581	17	Q9HLE2	Q9hle2 thermoplasm	216	33	51.6	483	13	Q9PMW7	Q9pmw7 xenopus lae
144	34	53.1	589	10	Q9LGC1	Q9lgc1 arabadopsi	217	33	51.6	521	16	Q980R8	Q980r8 mycoplasma
145	34	53.1	591	13	Q9PVX2	Q9pvx2 cynopus pyr	218	33	51.6	521	5	Q9GRG7	Q9grg7 trypanosoma
146	34	53.1	597	16	Q9RY11	Q9ry11 delnoccocu	219	33	51.6	547	16	Q9RZT4	Q9rzt4 delnoccocu
147	34	53.1	618	10	Q9LSH1	Q9lsh1 arabadopsi	220	33	51.6	557	12	Q9KR26	Q9kr26 bacillus ha
148	34	53.1	623	17	Q978S7	Q978s7 thermoplasm	221	33	51.6	562	12	Q9ORA4	Q9ora4 tomato ring
149	34	53.1	623	17	Q978S7	Q978s7 thermoplasm	222	33	51.6	564	11	Q9WUW1	Q9wuw1 mus musculu
150	34	53.1	683	5	Q17740	Q17740 caenorhabdi	223	33	51.6	572	11	Q99MX5	Q99mx5 cricetus
151	34	53.1	688	11	Q07643	Q07643 mus musculu	224	33	51.6	572	11	Q99MX5	Q99mx5 cricetus
152	34	53.1	689	16	Q8R808	Q8r808 thermomane	225	33	51.6	594	2	Q8ZFH4	Q8zfh4 yerishia pe
153	34	53.1	710	5	Q9XV52	Q9xv52 drosophila	226	33	51.6	613	13	Q9VSS5	Q9vss5 drosophila
154	34	53.1	710	5	Q9V78	Q9v78 drosophila	227	33	51.6	613	16	Q8RBV9	Q8rbv9 thermomane
155	34	53.1	710	5	Q8SAC5	Q8sac5 drosophila	228	33	51.6	628	10	Q9LP82	Q9lp82 arabadopsi
156	34	53.1	724	11	Q91YK2	Q91yk2 mus musculu	229	33	51.6	628	3	Q96X54	Q96x54 aspergillus
157	34	53.1	792	5	Q9VTV1	Q9vtv1 drosophila	230	33	51.6	650	16	Q8Y392	Q8y392 ralsstonia s
158	34	53.1	844	5	Q8SXX2	Q8sxx2 drosophila	231	33	51.6	743	4	Q8RB20	Q8rb20 homo sapien
159	34	53.1	881	16	Q9KDB8	Q9kdb8 bacillus ha	232	33	51.6	807	16	Q8XMR4	Q8xmr4 clostridium
160	34	53.1	910	16	Q8YGX8	Q8ygx8 bruceella me	233	33	51.6	866	10	Q40664	Q40664 oryza sativ
161	34	53.1	927	16	Q981B1	Q981b1 rhizobium l	234	33	51.6	926	13	Q93368	Q93368 apteronotus
162	34	53.1	938	4	Q96E80	Q96e80 homo sapien	235	33	51.6	966	16	Q92J03	Q92j03 tickettsia

236	33	51.6	972	16	P73273	P73273 synechocyst	309	32	50.0	533	16	Q8YND5	Q8YND5 anabaena sp
237	33	51.6	1011	3	Q9HEL5	Q9HEL5 neurospora	310	32	50.0	540	12	Q9IV49	Q9IV49 human calic
238	33	51.6	1225	3	Q9UJ98	Q9UJ98 homo sapien	311	32	50.0	540	12	Q9IBB0	Q9IBB0 norwalk-lik
239	33	51.6	1240	11	Q70576	Q70576 mus musculu	312	32	50.0	542	8	Q9BRU7	Q9BRU7 guillardi
240	33	51.6	1290	5	Q9NEA7	Q9NEA7 leishmania	313	32	50.0	545	3	P87306	P87306 schizosacch
241	33	51.6	1352	13	Q90YX0	Q90YX0 brachydanio	314	32	50.0	556	16	Q8VJ22	Q8VJ22 mycobacteri
242	33	51.6	1414	5	Q26634	Q26634 strongyloce	315	32	50.0	562	12	Q90RA5	Q90RA5 tomato ring
243	33	51.6	1675	12	Q90301	Q90301 rice tungro	316	32	50.0	572	5	Q20486	Q20486 caenorhabdl
244	33	51.6	1682	11	Q9QZK9	Q9QZK9 mus musculu	317	32	50.0	575	2	Q92P23	Q92P23 pseudomonas
245	33	51.6	1905	13	Q91823	Q91823 xenopus lae	318	32	50.0	576	2	Q82950	Q82950 pseudomonas
246	32.5	50.8	380	10	Q9S9V4	Q9S9V4 arabisopsis	319	32	50.0	592	6	Q9XSR3	Q9XSR3 canis famli
247	32.5	50.8	2471	3	Q59892	Q59892 candida alb	320	32	50.0	595	8	Q9M199	Q9M199 rhodospoete
248	32	50.0	77	9	Q9ZX19	Q9ZX19 bacterioph	321	32	50.0	595	8	Q9M199	Q9M199 typhlosecte
249	32	50.0	82	10	Q8RZD2	Q8RZD2 oryza sativ	322	32	50.0	605	16	Q92E44	Q92E44 listeria in
250	32	50.0	103	3	Q9S336	Q9S336 prochloroco	323	32	50.0	620	2	Q9ANF0	Q9ANF0 bradyrhizob
251	32	50.0	111	17	Q8ZXK0	Q8ZXK0 pyrobaculum	324	32	50.0	622	4	Q96C38	Q96C38 homo sapien
252	32	50.0	118	2	Q9Z5E4	Q9Z5E4 phytoplasma	325	32	50.0	631	10	Q9LZU1	Q9LZU1 arabisopsis
253	32	50.0	127	2	Q9AGM6	Q9AGM6 nostoc punc	326	32	50.0	641	11	Q922K2	Q922K2 mus musculu
254	32	50.0	134	12	Q99EM5	Q99EM5 human calic	327	32	50.0	771	5	Q19380	Q19380 caenorhabdl
255	32	50.0	175	17	Q9YD99	Q9YD99 aeropyrum p	328	32	50.0	803	2	Q93H55	Q93H55 bradyrhizob
256	32	50.0	186	13	Q9PSX5	Q9PSX5 paralicthy	329	32	50.0	806	2	Q45278	Q45278 bradyrhizob
257	32	50.0	187	11	Q9C0V9	Q9C0V9 mus musculu	330	32	50.0	812	5	Q06452	Q06452 ephydaria m
258	32	50.0	205	4	Q9H1F7	Q9H1F7 homo sapien	331	32	50.0	834	10	Q9RKE9	Q9RKE9 arabisopsis
259	32	50.0	211	10	Q9S1P2	Q9S1P2 carica papa	332	32	50.0	872	2	Q47114	Q47114 escherichia
260	32	50.0	214	3	Q9HEZ5	Q9HEZ5 neurospora	333	32	50.0	873	4	Q9UMF9	Q9UMF9 homo sapien
261	32	50.0	220	16	Q9KQ08	Q9KQ08 vibrio chol	334	32	50.0	873	4	Q96M72	Q96M72 homo sapien
262	32	50.0	221	16	Q97H92	Q97H92 clostridium	335	32	50.0	896	4	Q96SFO	Q96SFO homo sapien
263	32	50.0	221	16	Q9RUC5	Q9RUC5 streptomyce	336	32	50.0	901	13	Q921B1	Q921B1 mus musculu
264	32	50.0	243	5	Q9BKN6	Q9BKN6 strongyloce	337	32	50.0	901	13	Q91948	Q91948 xenopus sp.
265	32	50.0	247	5	Q9BKN4	Q9BKN4 strongyloce	338	32	50.0	901	16	Q9PH65	Q9PH65 xylella fas
266	32	50.0	247	5	Q9BKN5	Q9BKN5 strongyloce	339	32	50.0	902	10	Q9M364	Q9M364 arabisopsis
267	32	50.0	249	16	Q9BET7	Q9BET7 rhizobium 1	340	32	50.0	904	13	Q91977	Q91977 xenopus lae
268	32	50.0	250	5	Q9BKN4	Q9BKN4 strongyloce	341	32	50.0	1040	5	Q925I8	Q925I8 mus musculu
269	32	50.0	252	7	Q31369	Q31369 brachydanio	342	32	50.0	1060	5	Q9VSR1	Q9VSR1 drosophila
270	32	50.0	268	5	Q17030	Q17030 anopheles g	343	32	50.0	1064	10	Q04135	Q04135 arabisopsis
271	32	50.0	276	16	Q8XP0	Q8XP0 talstonia s	344	32	50.0	1064	10	Q04156	Q04156 arabisopsis
272	32	50.0	277	10	Q94L68	Q94L68 oryza sativ	345	32	50.0	1064	10	Q04884	Q04884 arabisopsis
273	32	50.0	293	17	Q9YDC3	Q9YDC3 aeropyrum p	346	32	50.0	1064	10	Q9SMZ4	Q9SMZ4 arabisopsis
274	32	50.0	297	16	Q8XCR2	Q8XCR2 escherichia	347	32	50.0	1093	4	Q9P2E7	Q9P2E7 homo sapien
275	32	50.0	304	16	Q91798	Q91798 pseudomonas	348	32	50.0	1129	5	Q93342	Q93342 caenorhabdl
276	32	50.0	306	16	Q8U5H3	Q8U5H3 agrobacteri	349	32	50.0	1140	11	Q61434	Q61434 mus musculu
277	32	50.0	307	5	Q18327	Q18327 caenorhabdl	350	32	50.0	1198	16	Q9CNS4	Q9CNS4 pasteurella
278	32	50.0	308	16	Q9X7E8	Q9X7E8 mycobacteri	351	32	50.0	1256	11	Q99M76	Q99M76 rattus norv
279	32	50.0	309	8	Q03140	Q03140 acorus gram	352	32	50.0	1264	12	Q89230	Q89230 variola vir
280	32	50.0	309	16	Q92UG2	Q92UG2 rhizobium m	353	32	50.0	1336	10	Q81074	Q81074 arabisopsis
281	32	50.0	310	16	Q9S2J4	Q9S2J4 streptomyce	354	32	50.0	1347	4	Q960B3	Q960B3 homo sapien
282	32	50.0	311	2	Q30422	Q30422 caldocellum	355	32	50.0	1378	5	Q97405	Q97405 halloia di
283	32	50.0	321	12	Q9J454	Q9J454 human rotav	356	32	50.0	1491	13	Q91718	Q91718 xenopus lae
284	32	50.0	330	10	Q946D5	Q946D5 arabisopsis	357	32	50.0	1497	11	Q61431	Q61431 mus musculu
285	32	50.0	333	16	Q9HMI3	Q9HMI3 pseudomonas	358	32	50.0	1524	10	Q8W5F5	Q8W5F5 oryza sativ
286	32	50.0	345	16	Q9CNK9	Q9CNK9 pasteurella	359	32	50.0	1551	5	Q9VMP3	Q9VMP3 drosophila
287	32	50.0	361	3	Q9UT98	Q9UT98 schizosacch	360	32	50.0	1573	5	Q95V18	Q95V18 drosophila
288	32	50.0	370	3	P78821	P78821 schizosacch	361	32	50.0	1648	11	Q9ESB7	Q9ESB7 mus musculu
289	32	50.0	376	11	Q9JK15	Q9JK15 rattus norv	362	32	50.0	1648	11	Q9Q205	Q9Q205 rice tungro
290	32	50.0	380	5	Q07138	Q07138 microclonia	363	32	50.0	1674	12	Q91D73	Q91D73 white spot
291	32	50.0	381	10	Q9LD25	Q9LD25 arabisopsis	364	32	50.0	1684	12	Q91LK9	Q91LK9 white spot
292	32	50.0	382	5	Q9N3T2	Q9N3T2 caenorhabdl	365	32	50.0	1701	3	Q8WZX8	Q8WZX8 white spot
293	32	50.0	390	10	Q9LJZ4	Q9LJZ4 oryza sativ	366	32	50.0	1774	11	Q62001	Q62001 mus musculu
294	32	50.0	399	16	Q98IP0	Q98IP0 rhizobium 1	367	32	50.0	1869	12	Q8V2H2	Q8V2H2 camelpox vi
295	32	50.0	403	16	Q8TEU7	Q8TEU7 homo sapien	368	32	50.0	1869	12	Q8QP27	Q8QP27 camelpox vi
296	32	50.0	403	16	Q99WH0	Q99WH0 staphylococ	369	32	50.0	1869	12	Q8QP27	Q8QP27 camelpox vi
297	32	50.0	421	16	Q84091	Q84091 chlamydia t	370	32	50.0	1896	12	Q89117	Q89117 variola vir
298	32	50.0	425	2	Q9RMB0	Q9RMB0 bacillus ce	371	32	50.0	1896	12	Q89117	Q89117 variola vir
299	32	50.0	443	3	Q9HGL3	Q9HGL3 schizosacch	372	32	50.0	1896	12	Q89096	Q89096 variola vir
300	32	50.0	443	4	Q9Y2B1	Q9Y2B1 homo sapien	373	32	50.0	1897	12	Q89096	Q89096 variola vir
301	32	50.0	448	10	Q9MAT6	Q9MAT6 arabisopsis	374	32	50.0	1919	12	Q8O4M6	Q8O4M6 variola maj
302	32	50.0	461	10	Q9FTD4	Q9FTD4 oryza sativ	375	32	50.0	1924	12	Q9JFS2	Q9JFS2 cowpox viru
303	32	50.0	467	16	Q8YWE0	Q8YWE0 anabaena sp	376	32	50.0	1933	12	Q72759	Q72759 ecrtomella
304	32	50.0	468	16	Q9L0P9	Q9L0P9 streptomyce	377	32	50.0	1958	11	P70276	P70276 cowpox viru
305	32	50.0	478	17	Q28457	Q28457 archaeoglob	378	32	50.0	2090	3	Q9P6X4	Q9P6X4 mus musculu
306	32	50.0	487	16	Q9ATV2	Q9ATV2 caulobacter	379	32	50.0	2684	10	Q9LNF9	Q9LNF9 neurospora
307	32	50.0	492	15	Q88987	Q88987 visna virus	380	32	50.0	4099	10	Q9C7Z6	Q9C7Z6 arabisopsis
308	32	50.0	527	16	Q97H21	Q97H21 clostridium	381	32	50.0	6239	2	Q9S0R7	Q9S0R7 streptomyce

382	31.5	49.2	580	16	Q8XRM3	Q8XRM3 ralslonia s	455	31	48.4	255	5	Q9BKM9	Q9BKM9 strongyloce	
383	31.5	49.2	632	17	Q9V229	Q9V229 pyrococcus	456	31	48.4	256	5	Q9BKN1	Q9BKN1 strongyloce	
384	31.5	49.2	641	17	Q8RTN5	Q8RTN5 pyrococcus	457	31	48.4	256	11	Q923P1	Q923P1 mus muscul	
385	31.5	49.2	788	10	Q94ZM4	Q94ZM4 oryza sativ	458	31	48.4	259	10	Q9SEW5	Q9SEW5 vicia faba	
386	31.5	49.2	819	11	Q923X3	Q923X3 cricetus	459	31	48.4	260	16	Q06345	Q06345 mycobacteri	
387	31.5	49.2	1900	12	Q9WTF1	Q9WTF1 sindbis-11k	460	31	48.4	262	5	Q9BKN0	Q9BKN0 strongyloce	
388	31.5	49.2	1901	12	Q884Z9	Q884Z9 sindbis-11k	461	31	48.4	263	10	Q9SEW7	Q9SEW7 vicia faba	
389	31.5	49.2	2500	12	Q884J1	Q884J1 sindbis-11k	462	31	48.4	266	5	P91608	Q9SEW7 vicia faba	
390	31.5	49.2	2513	12	Q87644	Q87644 sindbis vir	463	31	48.4	267	10	Q9SCR9	Q9SCR9 aradidopsi	
391	31.5	49.2	2517	12	Q9YJX6	Q9YJX6 sindbis-11k	464	31	48.4	269	16	Q8X613	Q8X613 escherichia	
392	31	48.4	51	16	Q8VWC2	Q8VWC2 mycobacteri	465	31	48.4	270	4	Q9NS61	Q9NS61 homo sapien	
393	31	48.4	79	9	Q9XJF0	Q9XJF0 lactococcus	466	31	48.4	270	4	Q9ET41	Q9ET41 homo sapien	
394	31	48.4	79	9	Q94M89	Q94M89 lactococcus	467	31	48.4	270	5	Q9BKN2	Q9BKN2 strongyloce	
395	31	48.4	81	2	Q9E0Z3	Q9E0Z3 escherichia	468	31	48.4	270	5	Q9T206	Q9T206 strongyloce	
396	31	48.4	98	8	Q99955	Q99955 cyprinella	469	31	48.4	270	5	Q96536	Q96536 strongyloce	
397	31	48.4	99	11	Q91WC4	Q91WC4 mus muscul	470	31	48.4	270	6	Q8WN03	Q8WN03 mustela put	
398	31	48.4	108	16	Q8Z6L4	Q8Z6L4 salmonella	471	31	48.4	270	11	Q9JMS9	Q9JMS9 rattus norv	
399	31	48.4	108	16	Q84944	Q84944 salmonella	472	31	48.4	270	11	Q9JUF6	Q9JUF6 mus muscul	
400	31	48.4	116	4	Q8TDC8	Q8TDC8 homo sapien	473	31	48.4	270	11	Q9J123	Q9J123 rattus norv	
401	31	48.4	134	16	Q92KX5	Q92KX5 rhizobium m	474	31	48.4	270	11	Q8VHN6	Q8VHN6 mus muscul	
402	31	48.4	136	11	Q923N9	Q923N9 mus muscul	475	31	48.4	272	10	Q943Z4	Q943Z4 aradidopsi	
403	31	48.4	139	9	Q9T1L4	Q9T1L4 bacterioph	476	31	48.4	277	5	Q9BKN3	Q9BKN3 strongyloce	
404	31	48.4	139	16	Q8X697	Q8X697 escherichia	477	31	48.4	283	16	Q9RX60	Q9RX60 delnoccoc	
405	31	48.4	146	16	Q92PML	Q92PML rhizobium m	478	31	48.4	285	4	Q96T44	Q96T44 homo sapien	
406	31	48.4	150	6	Q77799	Q77799 bos taurus	479	31	48.4	289	16	Q8Y1T5	Q8Y1T5 talstonia s	
407	31	48.4	155	11	Q9D9L7	Q9D9L7 mus muscul	480	31	48.4	290	2	Q8RLR1	Q8RLR1 acarcus to	
408	31	48.4	157	2	Q9X4L1	Q9X4L1 paracoccus	481	31	48.4	309	2	Q9F694	Q9F694 streptococ	
409	31	48.4	157	2	P95314	P95314 mycobacteri	482	31	48.4	311	16	Q9A3C4	Q9A3C4 caulobacter	
410	31	48.4	159	4	Q8WMD6	Q8WMD6 homo sapien	483	31	48.4	312	2	Q9F693	Q9F693 streptococ	
411	31	48.4	161	10	Q8WVU9	Q8WVU9 narcissus p	484	31	48.4	313	5	Q77284	Q77284 drosophila	
412	31	48.4	163	16	Q9HYA1	Q9HYA1 pseudomona	485	31	48.4	315	16	Q8XU25	Q8XU25 arabidopsi	
413	31	48.4	173	4	Q96CZ6	Q96CZ6 homo sapien	486	31	48.4	317	16	Q9KM71	Q9KM71 vldrio chol	
414	31	48.4	184	17	Q96T43	Q96T43 homo sapien	487	31	48.4	317	16	Q8YPM2	Q8YPM2 arabidopsi	
415	31	48.4	184	16	Q9YEB5	Q9YEB5 aeropyrum p	488	31	48.4	319	12	Q8Z7A3	Q8Z7A3 pyrobaculu	
416	31	48.4	188	16	Q97P28	Q97P28 streptococ	489	31	48.4	319	17	Q8Z7A3	Q8Z7A3 pyrobaculu	
417	31	48.4	192	16	Q8VJUR4	Q8VJUR4 mycobacteri	490	31	48.4	330	16	Q9EMF0	Q9EMF0 methanosarc	
418	31	48.4	197	16	Q9HT85	Q9HT85 pseudomona	491	31	48.4	333	17	Q8TUB9	Q8TUB9 human immun	
419	31	48.4	198	2	Q9KMN1	Q9KMN1 staphylococ	492	31	48.4	334	15	Q9TUB9	Q9TUB9 human immun	
420	31	48.4	206	2	Q925Y8	Q925Y8 corynebacte	493	31	48.4	339	10	Q9SK61	Q9SK61 aradidopsi	
421	31	48.4	208	4	Q9U777	Q9U777 homo sapien	494	31	48.4	342	2	Q8RT50	Q8RT50 pseudomona	
422	31	48.4	208	4	Q961Y6	Q961Y6 homo sapien	495	31	48.4	342	16	Q9AAV9	Q9AAV9 pseudomona	
423	31	48.4	209	10	Q23485	Q23485 aradidopsi	496	31	48.4	343	16	Q92ZV7	Q92ZV7 rhizobium m	
424	31	48.4	213	4	Q9UBM0	Q9UBM0 homo sapien	497	31	48.4	345	5	Q95YF7	Q95YF7 leishmania	
425	31	48.4	218	16	Q9KXY6	Q9KXY6 streptococ	498	31	48.4	346	5	Q96FM2	Q96FM2 ceenorhdbd1	
426	31	48.4	219	5	Q96537	Q96537 allocentrot	499	31	48.4	348	16	Q9KHM4	Q9KHM4 streptococ	
427	31	48.4	219	5	Q9TYD4	Q9TYD4 allocentrot	500	31	48.4	349	10	Q9SM53	Q9SM53 aradidopsi	
428	31	48.4	220	4	Q9NS60	Q9NS60 homo sapien								
429	31	48.4	220	4	Q9HD11	Q9HD11 homo sapien								
430	31	48.4	220	6	Q8WN04	Q8WN04 mustela put								
431	31	48.4	220	11	Q9JMF0	Q9JMF0 rattus norv								
432	31	48.4	220	11	Q9JMF0	Q9JMF0 rattus norv								
433	31	48.4	220	11	Q9JMF0	Q9JMF0 rattus norv								
434	31	48.4	220	11	Q9JMF0	Q9JMF0 rattus norv								
435	31	48.4	221	5	Q9S519	Q9S519 mus muscul								
436	31	48.4	221	5	Q9S519	Q9S519 mus muscul								
437	31	48.4	225	4	Q9H0N4	Q9H0N4 homo sapien								
438	31	48.4	225	4	Q96KR6	Q96KR6 homo sapien								
439	31	48.4	227	4	Q9HD10	Q9HD10 homo sapien								
440	31	48.4	228	2	Q9WZ44	Q9WZ44 pseudomona								
441	31	48.4	232	5	Q95363	Q95363 strongyloce								
442	31	48.4	232	5	Q25367	Q25367 loligo opal								
443	31	48.4	235	5	Q96535	Q96535 strongyloce								
444	31	48.4	236	5	Q96538	Q96538 strongyloce								
445	31	48.4	236	5	Q97Y05	Q97Y05 hemocentrot								
446	31	48.4	237	16	Q9AHJ5	Q9AHJ5 escherichia								
447	31	48.4	242	2	Q9E290	Q9E290 lactococcus								
448	31	48.4	243	10	Q8RX79	Q8RX79 aradidopsi								
449	31	48.4	245	2	Q93TM9	Q93TM9 nostoc punc								
450	31	48.4	252	4	Q9NZ11	Q9NZ11 homo sapien								
451	31	48.4	252	5	Q9BKM8	Q9BKM8 strongyloce								
452	31	48.4	252	6	Q8WN05	Q8WN05 mustela put								
453	31	48.4	252	11	Q9J122	Q9J122 rattus norv								
454	31	48.4	252	11	Q8VHN5	Q8VHN5 mus muscul								

AC	Q9R149;	PRELIMINARY;	PRT;	230 AA.	Q9BKM9	Q9BKM9 strongyloce
DT	01-MAY-2000 (TREMblrel. 13, Created)				Q9BKN1	Q9BKN1 strongyloce
DT	01-MAY-2000 (TREMblrel. 13, Last sequence update)				Q923P1	Q923P1 mus muscul
DT	01-JUN-2002 (TREMblrel. 21, Last annotation update)				Q9SEW5	Q9SEW5 vicia faba
DE	Pro-alpha-1 type I collagen (Fragment).				Q06345	Q06345 mycobacteri
OC	Cavia porcellus (Guinea pig).				Q9BKN0	Q9BKN0 strongyloce
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				Q9SEW7	Q9SEW7 vicia faba
OC	Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.				P91608	Q9SEW7 vicia faba
NCBI_TaxID=10141;						
[1]						
RP	SEQUENCE FROM N.A.					
RC	STRAIN-HARLEY; TISSUE=TRACHEA;					
RA	Morishima Y., Uchida Y., Nomura A., Ishii Y., Sakamoto T.,					
RA	Sekizawa K.;					
RT	"Guinea-pig pro-alpha-1 type I collagen expression in injured tracheal					
RT	epithelium".					
RT	Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.					
DR	EMBL; AF169346; AAD49346.1; -					
DR	InterPro; IPR000087; Collagen.					

ALIGNMENTS

RESULT 1

Q9R149 PRELIMINARY; PRT; 230 AA.

AC Q9R149;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 21, Last annotation update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Pro-alpha-1 type I collagen (Fragment).

OS Cavia porcellus (Guinea pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Hystriolognathi; Cavidae; Cavia.

OX NCBI_TaxID=10141;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=HARLEY; TISSUE=TRACHEA;

RA Morishima Y., Uchida Y., Nomura A., Ishii Y., Sakamoto T.,

RA Sekizawa K.;

RT "Guinea pig" pro-alpha-1 type I collagen expression in injured tracheal

RT epithelium.;

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF169346; AAD49346.1; -

InterPro; IPR000087; Collagen.

DR Pfam: PF01391; Collagen; 4.
 KW Collagen.
 FT NON_TER 1
 FT NON_TER 230
 SQ SEQUENCE 230 AA; 20425 MW; 1A465F92779D9A71 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 42; DB 11; Length 230;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 POGIAGOR 9
 DB 82 POGIAGOR 89

RESULT 2

O99LL6 PRELIMINARY; PRT; 589 AA.

AC O99LL6; 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DE Hypothetical 58.8 kDa protein (Fragment).
 GN COL1A1.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;

RN [1]
 RP SEQUENCE FROM N.A.

RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: BC003198; AAH03198.1;
 DR MGD: MGI:88467; Col1a1.

DR InterPro: IPR000087; Collagen.

DR InterPro: IPR000085; Fib_collagen_C.

DR Pfam: PF01410; COLF1; 1.

DR Pfam: PF01391; Collagen; 6.

DR Prodom: PD002078; Fib_collagen_C; 1.

DR SMART: SM00038; COLF1; 1.

DR Collagen: Hypothetical protein.

FT NON_TER 1
 SO SEQUENCE 589 AA; 58805 MW; 81847495E505CEF CRC64;

Query Match
 Best Local Similarity 100.0%; Score 42; DB 11; Length 589;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 POGIAGOR 9
 DB 76 POGIAGOR 83

RESULT 3

O9YIB4 PRELIMINARY; PRT; 1450 AA.

AC O9YIB4; 01-MAY-1999 (TREMBLrel. 10, Created)

DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Alpha 1 type I collagen.

OS Cynops pyrrhogaster (Japanese common newt).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Caudata; Salamandridae; Cynops.

OX NCBI_TaxID=8330;

RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE-REGNERATE FORELIMS;

RX MEDLINE-99407244; PubMed-10474166;

RA Asahina K., Obara M., Yoshizato K.;

RT "Expression of genes of type I and type II collagen in the formation
 and development of the blastema of regenerating newt limb.";
 Dev. Dyn. 216:59-71(1999).

DR EMBL: AB015438; BAA36973.1; -
 DR InterPro: IPR000087; Collagen.
 DR InterPro: IPR000085; Fib_collagen_C.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF01410; COLF1; 1.
 DR Pfam: PF01391; Collagen; 18.
 DR Prodom: PD000007; Collagen; 1.
 DR Prodom: PD002078; Fib_collagen_C; 1.
 DR SMART: SM00038; COLF1; 1.
 DR SMART: SM00214; VWF; 1.
 DR PROSITE: PS01208; VWF; UNKNOWN_1.
 DR Collagen.
 SQ SEQUENCE 1450 AA; 137564 MW; ABF8A74841B8767C CRC64;

Query Match
 Best Local Similarity 100.0%; Score 42; DB 13; Length 1450;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 POGIAGOR 9
 DB 937 POGIAGOR 944

RESULT 4

O63079 PRELIMINARY; PRT; 1453 AA.

AC O63079; 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Collagen alpha1 (Fragment).

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]
 RP SEQUENCE OF 1-1092 FROM N.A.

RA STRAIN-SPRAGUE-DAWLEY; TISSUE-TOOTH;

RC Brandsten C., Lundmark C., Christenson C., Hammarstrom L., Wurtz T.;

RT "Expression of Collagen alpha1(I) mRNA variants during tooth and bone
 formation in the Rat.";

J. Dent. Res. 0:0-0(0).

FT NON_TER 1
 SO SEQUENCE 1453 AA; 137887 MW; E6896BDC19A4A1D8 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 42; DB 11; Length 1453;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 POGIAGOR 9
 DB 940 POGIAGOR 947

RESULT 5

O63079 PRELIMINARY; PRT; 1453 AA.

AC O63079; 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Collagen alpha1 (Fragment).

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]
 RP SEQUENCE OF 1-1092 FROM N.A.

RA STRAIN-SPRAGUE-DAWLEY; TISSUE-TOOTH;

RC Brandsten C., Lundmark C., Christenson C., Hammarstrom L., Wurtz T.;

RT "Expression of Collagen alpha1(I) mRNA variants during tooth and bone
 formation in the Rat.";

J. Dent. Res. 0:0-0(0).

FT NON_TER 1
 SO SEQUENCE 1453 AA; 137887 MW; E6896BDC19A4A1D8 CRC64;

ID 076045 PRELIMINARY; PRT; 1461 AA.
 AC 076045;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Pro alpha 1(I) collagen.
 GN COL1A1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN 111
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85130970; PubMed=2857713;
 RA Chu M.L., de Wet W., Bernard M., Ramirez F.;
 RT "Fine structural analysis of the human pro-alpha 1(I) collagen gene.
 RT Promoter structure, AluI repeats, and polymorphic transcripts.";
 RL J. Biol. Chem. 260:2315-2320(1985).
 RN 12
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88329734; PubMed=2843432;
 RA D'Alessio M., Bernard M., Pretorius P.J., de Wet W., Ramirez F.;
 RT "Complete nucleotide sequence of the region encompassing the first
 RT twenty-five exons of the human pro alpha 1(I) collagen gene.";
 RL Gene 67:105-115(1988).
 RN 13
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89025644; PubMed=3178743;
 RA Tromp G., Kuivaniemi H., Stacey A., Shikata H., Baldwin C.T.,
 RA Jaenisch R., Prockop D.J.;
 RT "Structure of a full-length cDNA clone for the prepro alpha 1(I) chain
 RT of human type I procollagen.";
 RL Biochem. J. 253:919-922(1988).
 RN 14
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91138770; PubMed=1995349;
 RA Maatta A., Bornstein P., Penttinen R.P.;
 RT "Highly conserved sequences in the 3'-untranslated region of the
 RT COL1A1 gene bind cell-specific nuclear proteins.";
 RL FEBS Lett. 279:9-13(1991).
 RN 15
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92157916; PubMed=1787829;
 RA Westerhausen A., Constantinou C.D., Pack M., Peng M.Z., Hanning C.,
 RA Olsen A.S., Prockop D.J.;
 RT "Completion of the last half of the structure of the human gene for
 RT the Pro alpha 1(I) chain of type I procollagen (COL1A1).";
 RL Matrix 11:375-379(1991).
 RN 16
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98107942; PubMed=9443882;
 RA Korhko J., Ala-Korhko L., De Paape A., Nuytlinck L., Earley J.,
 RA Prockop D.J.;
 RT "Analysis of the COL1A1 and COL1A2 genes by PCR amplification and
 RT scanning by conformation-sensitive gel electrophoresis identifies only
 RT COL1A1 mutations in 15 patients with osteogenesis imperfecta type I:
 RT Identification of common sequences of null-allele mutations.";
 RL Am. J. Hum. Genet. 62:98-110(1998).
 RN 17
 RP SEQUENCE FROM N.A.
 RA Korhko J.M., Earley J.J., Nuytlinck L., DePaape A., Prockop D.J.,
 RA Ala-Korhko L.;
 RN Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF017178; AAB94054.2; -;
 DR InterPro: IPR000087; Collagen.
 DR InterPro: IPR000885; Fib_collagen_C.
 DR InterPro: IPR001007; WVF_C.
 DR Pfam: PF01410; COLFI; 1.
 DR Pfam: PF01391; Collagen; 18.
 DR Pfam: PF00093; WVC; 1.
 DR ProDom: PD000007; Collagen; 2.
 DR ProDom: PD002078; Fib_collagen_C; 1.
 DR SMART: SM00038; COLFI; 1.

DR SMART: SM00214; WVC; 1.
 DR PROSITE: PS01208; WVFC; 1.
 KW Collagen.
 SQ SEQUENCE 1461 AA; 138630 MW; 9ACF6DE30EA78E21 CRC64;
 Query Match 65.6%; Score 42; DB 4; Length 1461;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 POGIACOR 9
 Db 948 POGIACOR 955
 RESULT 6
 ID 08V172 PRELIMINARY; PRT; 113 AA.
 AC 08V172;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Collagen type II (Fragment).
 OC Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.
 OX NCBI_TaxID=10141;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=HARTLEY.
 RA Huebner J.L., Clark A.E., Kraus V.B., Otterness I.G.;
 RT "Collagen type II in the guinea pig.";
 RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF152862; AAL55558.1; -;
 DR InterPro: IPR000087; Collagen.
 DR Pfam: PF01391; Collagen; 1.
 DR ProDom: PD000007; Collagen; 1.
 FT NON_TER 1 113
 FT NON_TER 1 113
 SQ SEQUENCE 113 AA; 10284 MW; F7861901127A9BCE CRC64;
 Query Match 62.5%; Score 40; DB 11; Length 113;
 Best Local Similarity 87.5%; Pred. No. 5.4;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 2 POGIACOR 9
 Db 93 POGIACOR 100
 RESULT 7
 ID 09XT25 PRELIMINARY; PRT; 347 AA.
 AC 09XT25;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Type II collagen cyanogen bromide fragment CB10 (Fragment).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN 11
 RP SEQUENCE FROM N.A.
 RC TISSUE=CARTRIDGE;
 RA Tang B., Chiang T.M., Brand D.D., Gumanovskaya M.L., Stuart J.M.,
 RA Kang A.H., Myers L.K.;
 RT "Molecular Definition and Characterization of Recombinant Bovine CB8
 RT And CB10: Immunogenicity and Arthritisogenicity.";
 RL J. Clin. Immunol. 0:0-0(1999).
 DR EMBL: AF138883; AAD42346.1; -;
 DR InterPro: IPR000087; Collagen.
 DR Pfam: PF01391; Collagen; 6.

DR ProDom: PD000007; Collagen; 2.
 KW Collagen.
 FT NON_TER 1 1
 SO SEQUENCE 347 AA; 3105 MM; 5D1IC0AF34089DF6 CRC64;

Query Match
 Best Local Similarity 62.5%; Score 40; DB 6; Length 347;
 Pred. No. 18;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 POGIAGOR 9
 DB 223 POGIAGOR 230

RESULT 8
 ID 080592 PRELIMINARY; PRT; 475 AA.
 AC 080592;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Putative amino acid perase.
 GN T2711.3
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxId=3702;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=CV. COLUMBIA;
 RA Federle N.A., Palm C.J., Conway A.B., Kurtz D.B., Conway A.R.,
 RA Au M., Araujo R., Buehler E., Dewar K., Feng J., Kim C., Li Y.,
 RA O.J.O., Osborne B.I., Shinn P., Sun H., Tortum M., Vysotskaia V.S.,
 RA Yu G., Ecker J., Theologis A., Davis R.W.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC004122; AAC34329.1;
 DR InterPro: IPR002422; AA/rel_primease2.
 DR Pfam: PF01490; Aa_trans.1;
 SO SEQUENCE 475 AA; 51815 MM; B6DE571E2FE7F92 CRC64;

Query Match
 Best Local Similarity 58.3%; Score 40; DB 10; Length 475;
 Pred. No. 25;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 SPOGIAGORFN 12
 DB 87 SPDSITGRN 98

RESULT 9
 ID 014046 PRELIMINARY; PRT; 1160 AA.
 AC 014046;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE COL2A1 protein precursor (Fragment).
 GN COL2A1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP TISSUE=CARTILAGE;
 RA MEDLINE-90026318; PubMed-2803268;
 RA Baldwin C.T., Reginato A.M., Smith C., Jimenez S.A., Prockop D.J.;
 RT "Structure of cDNA clones coding for human type II procollagen. The
 RT alpha 1(I) chain is more similar to the alpha 1(I) chain than two
 RT other alpha chains of fibrillar collagens.";
 RL Biochem. J. 262:521-528(1989).

DR EMBL: X16711; CAA34683.1; -
 DR InterPro: IPR000087; Collagen.
 DR Pfam: PF01391; Collagen; 18.
 DR ProDom: PD000007; Collagen; 4.
 KW Collagen; Signal.
 FT SIGNAL 1 23
 FT CHAIN 113 >1160
 FT NON_TER 1160 1160
 SO SEQUENCE 1160 AA; 105630 MM; A7F0523B856C8639 CRC64;

Query Match
 Best Local Similarity 62.5%; Score 40; DB 4; Length 1160;
 Pred. No. 63;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 POGIAGOR 9
 DB 904 POGIAGOR 911

RESULT 10
 ID 028396 PRELIMINARY; PRT; 1418 AA.
 AC 028396;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Type II collagen.
 OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OX NCBI_TaxId=9796;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP Richardson D.W., Dodge G.R.;
 RA Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RA MacLeod J.N., Fubini S.L., Gu D.N., Tetreault J.W., Todhunter R.J.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U62528; AAB05773.1;
 DR EMBL; AF040638; AAB96768.1;
 DR InterPro: IPR000087; Collagen.
 DR Pfam: PF01410; COLFI; 1.
 DR Pfam: PF01391; Collagen; 18.
 DR ProDom: PD000007; Collagen; 4.
 DR ProDom: PD002078; FIB-collagen_C; 1.
 DR SMART: SM00038; COLFI; 1.
 KW Collagen.
 SO SEQUENCE 1418 AA; 134343 MM; 115FCD19EB8696A3 CRC64;

Query Match
 Best Local Similarity 62.5%; Score 40; DB 6; Length 1418;
 Pred. No. 78;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 POGIAGOR 9
 DB 904 POGIAGOR 911

RESULT 11
 ID 09W789 PRELIMINARY; PRT; 1418 AA.
 AC 09W789;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Alpha type II collagen.
 OS Cynops pyrrhogaster (Japanese common newt).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Caudata; Salamandridae; Salamandridae; Cynops.
 OX NCBI_TaxId=8330;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=99407244; PubMed=10474166;
 RA Asahina K., Obara M., Yoshizato K.;
 RT "Expression of genes of type I and type II collagen in the formation
 and development of the blastema of regenerating newt limb.";
 RL Dev. Dyn. 216:59-71(1999).
 DR EMBL: AB022046; BA82043.1; -
 DR InterPro: IPR000087; Collagen.
 DR InterPro: IPR000885; Fib_collagen_C.
 DR Pfam: PF01410; COLFI; 1.
 DR Pfam: PF01391; Collagen; 18.
 DR ProDom: PD000007; Collagen; 7.
 DR ProDom: PD002078; Fib_collagen_C; 1.
 DR SMART: SM00038; COLFI; 1.
 KW Collagen.
 SQ SEQUENCE 1418 AA; 135067 MW; C19A6E601A2A717E CRC64;

Query Match 62.5%; Score 40; DB 13; Length 1418;
 Best Local Similarity 87.5%; Pred. No. 78;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 POGIAGOR 9
 DB 904 POGIAGOR 911

RESULT 12
 Q63123 PRELIMINARY; PRT; 1419 AA.
 AC Q63123; Q63565;
 DT 01-NOV-1996 (TRENBLREL. 01, Created)
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
 DE Collagen alpha 1 type II (T1 mRNA).
 GN T1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BONE FRACTURE CALLUS;
 RA Urahe K., Sarkar G., Bolander M.E.;
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1372-1419 FROM N.A.
 RA Wurtz T., Brandsten C., Lundmark C., Christersson C.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 370-422 FROM N.A.
 RC STRAIN-DA; TISSUE-CARTILAGE;
 RX MEDLINE=94321934; PubMed=8046350;
 RA Michelson E., Malmstrom V., Reis S., Engstrom A., Burkhardt H.,
 RA Holmdahl R.;
 RT "T cell recognition of carbohydrates on type II collagen.";
 RL J. Exp. Med. 180:745-749(1994).
 DR EMBL: I48440; AA79780.1; -
 DR EMBL: AJ224879; CA12179.1; -
 DR EMBL: X79816; CA56213.1; -
 DR InterPro: IPR000087; Collagen.
 DR InterPro: IPR000885; Fib_collagen_C.
 DR Pfam: PF01410; COLFI; 1.
 DR Pfam: PF01391; Collagen; 18.
 DR ProDom: PD000007; Collagen; 18.
 DR ProDom: PD002078; Fib_collagen_C; 1.
 DR SMART: SM00038; COLFI; 1.
 KW Collagen.
 SQ SEQUENCE 1419 AA; 134570 MW; B7C63B77819CE50B CRC64;

Query Match 62.5%; Score 40; DB 11; Length 1419;
 Best Local Similarity 87.5%; Pred. No. 78;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 POGIAGOR 9
 DB 905 POGIAGOR 912

RESULT 13
 Q90W37 PRELIMINARY; PRT; 1420 AA.
 AC Q90W37;
 DT 01-DEC-2001 (TRENBLREL. 19, Created)
 DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
 DE Alpha 1 type IIA collagen precursor.
 GN COL2A1.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OX Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=STERNAL;
 RA Caixia X., Yongzhi X., Siqi G., Yiyang S.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY046949; AAK98621.1; -
 DR InterPro: IPR000087; Collagen.
 DR InterPro: IPR000885; Fib_collagen_C.
 DR Pfam: PF01410; COLFI; 1.
 DR Pfam: PF01391; Collagen; 18.
 DR ProDom: PD000007; Collagen; 6.
 DR ProDom: PD002078; Fib_collagen_C; 1.
 KW Collagen; Signal.
 FT SIGNAL 1 25
 SQ SEQUENCE 1420 AA; 134999 MW; 88D9AAB17P214FF5 CRC64;

Query Match 62.5%; Score 40; DB 13; Length 1420;
 Best Local Similarity 87.5%; Pred. No. 78;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 POGIAGOR 9
 DB 906 POGIAGOR 913

RESULT 14
 Q62031 PRELIMINARY; PRT; 1442 AA.
 AC Q62031;
 DT 01-NOV-1996 (TRENBLREL. 01, Created)
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
 DE Pro-alpha 1 type II collagen.
 GN COL2A1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57/BLACK;
 RX MEDLINE=93358489; PubMed=1885613;
 RA Metzaranta M., Roman D., de Crombrughe B., Vuorio E.;
 RT "Mouse type II collagen gene. Complete nucleotide sequence, exon
 structure, and alternative splicing.";
 RL J. Biol. Chem. 266:16862-16869(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57/BLACK;
 RA Vuorio E.;
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: M65161; AA68099.1; -
 DR MGI: 88452; Col2a1.
 DR InterPro: IPR000087; Collagen.

DR InterPro: IPR000885; Fib_collagen_C.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF01410; COLFI; 1.
 DR Pfam: PF01391; Collagen; 17.
 DR Pfam: PF00093; VWC; 1.
 DR ProDom: PD000007; Collagen; 5.
 DR ProDom: PD002078; Fib_collagen_C; 1.
 DR SMART: SM00038; COLFI; 1.
 DR SMART: SM00214; VWC; 1.
 DR PROSITE: PS01208; VWF; 1.
 DR Collagen.
 SO SEQUENCE 1442 AA; 137911 MW; D3A3274493C8821C CRC64;

Query Match 62.5%; Score 40; DB 11; Length 1442;
 Best Local Similarity 87.5%; Pred. No. 80;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 POGIAGOR 9
 DB 928 POGIAGOR 935

RESULT 15

062033 PRELIMINARY; PRT; 1442 AA.
 ID 062033;
 AC 062033;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Pro-alpha-1 type II collagen.
 GN COL2A1
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57/BLACK;
 RX MEDLINE=91358489; PubMed=1885613;
 RA Metzaranta M., Toman D., de Crombrughe B., Vuorio E.;
 RT "Mouse type II collagen gene. Complete nucleotide sequence, exon
 structure, and alternative splicing."
 RT J. Biol. Chem. 266:16862-16869(1991).
 RL [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57/BLACK;
 RA Vuorio E.;
 RL Submitted (JAN-1995) to the EMBL/Genbank/DBJ databases.
 DR EMBL: M65161; AAA68102.1; -
 DR MGD: MGI:88452; Col2a1.
 DR InterPro: IPR000887; Collagen.
 DR InterPro: IPR000885; Fib_collagen_C.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF01410; COLFI; 1.
 DR Pfam: PF01391; Collagen; 17.
 DR Pfam: PF00093; VWC; 1.
 DR ProDom: PD000007; Collagen; 5.
 DR ProDom: PD002078; Fib_collagen_C; 1.
 DR SMART: SM00038; COLFI; 1.
 DR SMART: SM00214; VWC; 1.
 DR PROSITE: PS01208; VWF; 1.
 DR Collagen.
 SO SEQUENCE 1442 AA; 137829 MW; F0E77C11BCAFA93B CRC64;

Query Match 62.5%; Score 40; DB 11; Length 1442;
 Best Local Similarity 87.5%; Pred. No. 80;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 POGIAGOR 9
 DB 928 POGIAGOR 935

RESULT 16

062032 PRELIMINARY; PRT; 1459 AA.
 ID 062032;
 AC 062032;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Pro-alpha-1 type II collagen.
 GN COL2A1
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57/BLACK;
 RX MEDLINE=91358489; PubMed=1885613;
 RA Metzaranta M., Toman D., de Crombrughe B., Vuorio E.;
 RT "Mouse type II collagen gene. Complete nucleotide sequence, exon
 structure, and alternative splicing."
 RT J. Biol. Chem. 266:16862-16869(1991).
 RL [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57/BLACK;
 RA Vuorio E.;
 RL Submitted (JAN-1995) to the EMBL/Genbank/DBJ databases.
 DR EMBL: M65161; AAA68101.1; -
 DR MGD: MGI:88452; Col2a1.
 DR InterPro: IPR000887; Collagen.
 DR InterPro: IPR000885; Fib_collagen_C.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF01410; COLFI; 1.
 DR Pfam: PF01391; Collagen; 18.
 DR Pfam: PF00093; VWC; 1.
 DR ProDom: PD000007; Collagen; 5.
 DR ProDom: PD002078; Fib_collagen_C; 1.
 DR SMART: SM00038; COLFI; 1.
 DR SMART: SM00214; VWC; 1.
 DR PROSITE: PS01208; VWF; 1.
 DR Collagen.
 SO SEQUENCE 1459 AA; 139071 MW; A09D24BF7357C827 CRC64;

Query Match 62.5%; Score 40; DB 11; Length 1459;
 Best Local Similarity 87.5%; Pred. No. 81;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 POGIAGOR 9
 DB 945 POGIAGOR 952

RESULT 17

014047 PRELIMINARY; PRT; 1487 AA.
 ID 014047;
 AC 014047;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Alpha-1 type II collagen.
 GN COL2A1
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BL00D;
 RX MEDLINE=85190534; PubMed=3857598;
 RA Cheah K.S., Stoker N.G., Griffin J.R., Grosveld F.G., Solomon E.;
 RT "Identification and characterization of the human type II collagen
 gene (COL2A1)."
 RT Proc. Natl. Acad. Sci. U.S.A. 82:2555-2559(1985).
 RL [2]

RP SEQUENCE FROM N.A.
 RC TISSUE-BLOOD:
 RX MEDLINE=90026318; PubMed=2803268;
 RA Baldwin C.T., Reginato A.M., Smith C., Jimenez S.A., Prockop D.J.;
 RT "Structure of cDNA clones coding for human type II procollagen. The
 RT alpha 1(I) chain is more similar to the alpha 1(I) chain than two
 RT other alpha chains of fibrillar collagens."; Biochem. J. 262:521-528(1989).
 RL Biochem. J. 262:521-528(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BLOOD:
 RX MEDLINE=89325561; PubMed=2753125;
 RA Vlkula M., Peltonen L.;
 RT "Structural analyses of the polymorphic area in type II collagen
 RT gene."; FEBS Lett. 250:171-174(1989).
 RL FEBS Lett. 250:171-174(1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BLOOD:
 RX MEDLINE=91184811; PubMed=2081599;
 RA Ryan M.C., Sieraski M., Sandell L.J.;
 RT "The human type II procollagen gene: identification of an additional
 RT protein-coding domain and location of potential regulatory sequences
 RT in the promoter and first intron."; Genomics 8:41-48(1990).
 RL Genomics 8:41-48(1990).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BLOOD:
 RX MEDLINE=91153296; PubMed=1999183;
 RA Huang M.C., Seyer J.M., Thompson J.P., Spinella D.G., Cheah K.S.,
 RT "Genomic organization of the human procollagen alpha 1(II) collagen
 RT gene."; Eur. J. Biochem. 195:593-600(1991).
 RL Eur. J. Biochem. 195:593-600(1991).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BLOOD:
 RX MEDLINE=92344585;
 RA Vlkula M., Metsaranta M., Syvanen A.C., Ala-Kokko L., Vuorio E.,
 RT "Structural analysis of the regulatory elements of the type-II
 RT procollagen gene. Conservation of promoter and first intron sequences
 RT between human and mouse."; Biochem. J. 285:0-0(0).
 RL Biochem. J. 285:0-0(0).
 RN [7]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BLOOD:
 RX MEDLINE=97104294; PubMed=8948452;
 RA Ala-Kokko L., Kvist A.P., Metsaranta M., Kivirikko K.I.,
 RT "Conservation of the sizes of 5' introns and over 100 intronic
 RT sequences for the binding of common transcription factors in the human
 RT and mouse genes for type II procollagen (COL2A1)."; EMBL. L10347; AAC41772.1; -.
 RL EMBL. L10347; AAC41772.1; -.
 DR InterPro: IPR000087; Collagen.
 DR InterPro: IPR000885; Fib_collagen_C.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF01410; COLFI; 1.
 DR Pfam: PF01391; Collagen; 18.
 DR Pfam: PF00093; VWC; 1.
 DR ProDom: PD000007; Collagen; 4.
 DR ProDom: PD002078; Fib_collagen_C; 1.
 DR SMART: SM00038; COLFI; 1.
 DR SMART: SM00214; VWC; 1.
 DR PROSITE: PS01208; VWF; 1.
 KM Collagen.
 SQ SEQUENCE 1487 AA; 141772 MW; 0B7E79D46BDAFA97 CRC64;

Query Match 62.5%; Score 40; DB 4; Length 1487;
 Best Local Similarity 87.5%; Pred. No. 82;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 POGIAGOR 9
 DB 973 POGIAGOR 980
 RESULT 18
 ID 077753 PRELIMINARY; PRT; 1487 AA.
 AC 077753;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Type IIA procollagen.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96340920; PubMed=9676231;
 RA Du F., Acland G.M., Ray J.;
 RT "Differential splicing of type II procollagen mRNA in canine retina.";
 RL Anim. Biotechnol. 9:15-20(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20480698; PubMed=11024291;
 RA Du F., Acland G.M., Ray J.;
 RT "Cloning and expression of type II collagen mRNA: evaluation as a
 RT candidate for canine ocular skeletal dysplasia."; Gene 255:307-316(2000).
 RL Gene 255:307-316(2000).
 DR EMBL: AF023169; AAC62178.2; -.
 DR InterPro: IPR00087; Collagen.
 DR InterPro: IPR000885; Fib_collagen_C.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF01410; COLFI; 1.
 DR Pfam: PF01391; Collagen; 18.
 DR Pfam: PF00093; VWC; 1.
 DR ProDom: PD000007; Collagen; 4.
 DR ProDom: PD002078; Fib_collagen_C; 1.
 DR SMART: SM00038; COLFI; 1.
 DR SMART: SM00214; VWC; 1.
 DR PROSITE: PS01208; VWF; 1.
 KM Collagen.
 SQ SEQUENCE 1487 AA; 141875 MW; 25873EAC1311DB8 CRC64;

Query Match 62.5%; Score 40; DB 6; Length 1487;
 Best Local Similarity 87.5%; Pred. No. 82;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 POGIAGOR 9
 DB 973 POGIAGOR 980
 RESULT 19
 ID 09KM17 PRELIMINARY; PRT; 211 AA.
 AC 09KM17;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE DNA insertion sequence IS1421 (Fragment).
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
 OC Ralstonia.
 OX NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF 301556;
 RA Hasebe A.;
 RT "Identification and nucleotide sequence of Ralstonia solanacearum
 RT insertion sequence IS1421, a new member of the IS427 subgroup of the
 RT ISS family.";

RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB045354; BAA97978.1; -
 DR InterPro: IPR002559; Transposase_11.
 DR Pfam: PF01609; Transposase_11; 1.
 FT NON_TER
 SQ SEQUENCE 211 AA; 23469 MW; 1AA5216B25677A8 CRC64;

Query Match
 Best Local Similarity 77.8%; Score 39; DB 2; Length 211;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 SPGINGOR 9
 DB 13 SPAGVAGOR 21

RESULT 20

ID 09KHH7 PRELIMINARY; PRT; 373 AA.

AC 09KHH7;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE YSAW.

OS YSAW.

OC Yersinia enterocolitica.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Yersinia.

OC NCBI_TaxID=630;

OX NCBI_TaxID=630;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE-20402121; PubMed-10931293;

RA Haller J.C., Carlson S., Pederson K.J., Piersen D.E.;

RT "A chromosomally-encoded type III secretion pathway in Yersinia

RT enterocolitica is important in virulence.";

RL Mol. Microbiol. 36:1436-1446(2000).

RN [2]

RP SEQUENCE FROM N.A.

RA Carlson S., Piersen D.E.;

RT Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA Haller J.C., Carlson S., Pederson K.J., Piersen D.E.;

RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF005744; AAF82328.1; -

DR InterPro: IPR003520; Invas_inve.

DR Pfam: PF02523; Inve; 1

SO SEQUENCE 373 AA; 41968 MW; 898EE5979A9DF3ED CRC64;

Query Match
 Best Local Similarity 59.4%; Score 38; DB 2; Length 373;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 GINGORNFN 12
 DB 287 GIRGOENFN 295

RESULT 21

ID 093NK8 PRELIMINARY; PRT; 373 AA.

AC 093NK8;

DT 01-DEC-2001 (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE YSAW.

OS YSAW.

OC Yersinia enterocolitica.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Yersinia.

OX NCBI_TaxID=630;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-A127;
 RA Foulter B.G.F., Mueller S., Purnelle B., Troisfontaines P.,
 RA Cornu G.R., (APR-2001) to the EMBL/GenBank/DBJ databases.
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF369544; AAK84112.1; -
 DR InterPro: IPR003520; Invas_inve.
 DR Pfam: PF02523; Inve; 1.
 SQ SEQUENCE 373 AA; 41968 MW; AFEF06F9863E904A CRC64;

Query Match
 Best Local Similarity 59.4%; Score 38; DB 2; Length 373;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 GINGORNFN 12
 DB 287 GIRGOENFN 295

RESULT 22

ID 082804 PRELIMINARY; PRT; 695 AA.

AC 082804;

DT 01-NOV-1998 (TREMBlrel. 08, Created)

DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE AT325930 protein (Hypothetical 77.2 kDa protein).

GN AT325930

OS Arabidopsis thaliana (mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OC NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN-CV, COLUMBIA;

RT MEDLINE-20083487; PubMed-10617197;

RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,

RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,

RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,

RA Cronin L.A., Shen M., Vanden S.E., Umayam L., Tallon L.J., Gill J.E.,

RA Adams M.D., Carreira A.J., Creasy T.H., Goodman H.M., Somerville C.R.,

RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,

RA Salzberg S.L., Fraser C.M., Venter J.C.;

RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis

RT thaliana.";

RL Nature 402:761-768(1999).

RN [2]

RP SEQUENCE FROM N.A.

RA STRAIN-CV, COLUMBIA;

RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RA STRAIN-CV, COLUMBIA;

RT Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: AC004747; AAC31242.1; -

DR EMBL: AC005395; AAM15042.1; -

KW Hypothetical protein

SO SEQUENCE 695 AA; 77206 MW; 607A0720ED381C08 CRC64;

Query Match
 Best Local Similarity 59.4%; Score 38; DB 10; Length 695;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 POGINGORNF 11
 |||||:|:|

Db 608 POGISGSKSF 617

RESULT 23

ID 004419 PRELIMINARY; PRT: 702 AA.

AC 004419; (TREMBlrel. 04, Created)
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
 DE Nematode responsive protein.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eucots II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-SHOOT; PubMed=10571046;

RX MEDLINE=20035823; PubMed=10571046;

RA Puzio P.S., Lausen J., Almeida-Engler J., Cai D., Gheysen G.,

GRUNDLER G.M.W.;

RT "Isolation of a gene from Arabidopsis thaliana related to nematode

feeding structures.";

RL EMBL, Y11994; CAA72719.1; -

DR SEQUENCE 702 AA; 78184 MW; F52F426657817954 CRC64;

SQ

Query Match 59.4%; Score 38; DB 10; Length 702;

Best Local Similarity 60.0%; Pred. No. 91;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 POGIAGORNF 11

Db 608 POGISGSKSF 617

RESULT 24

ID 093485 PRELIMINARY; PRT: 809 AA.

AC 093485; (TREMBlrel. 08, Created)

DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Alpha 1 type I collagen (Fragment).

OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Actinopterygii; Neopterygii; Teleostei; Euteleostei;

CC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

NCBI_TaxID=8022;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-FIBROBLAST;

RA Saito M., Kunisaki N., Hiroto I., Aoki T., Ishida M., Urano N.,

Kimura S.;

RT "Partial characterization of cDNA clones encoding the three distinct

pro alpha chains of type I collagen from rainbow trout.";

RL Fisheries Sci. 64:780-786(1998).

DR EMBL; AB008373; BAA33380.1; -

DR InterPro: IPR000087; Collagen.

DR Pfam: PF01410; COLFI.1.

DR Pfam: PF01391; Collagen.9.

DR ProDom: PD000007; Collagen.1.

DR ProDom: PD002078; Fib_collagen_C.1.

DR SMART; SM00038; COLFI.1.

KM Collagen.

FT NON TER

SQ SEQUENCE 809 AA; 78164 MW; 68C056A7640FC81 CRC64;

Query Match 59.4%; Score 38; DB 13; Length 809;

Best Local Similarity 87.5%; Pred. No. 11e+02;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 POGIAGOR 9

Db 296 POGIGCOR 303

RESULT 25

ID 0910C0 PRELIMINARY; PRT: 1449 AA.

AC 0910C0; (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Collagen a1(I).

OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Actinopterygii; Neopterygii; Teleostei; Euteleostei;

CC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.

NCBI_TaxID=8022;

RN [1]

RP SEQUENCE FROM N.A.

RC MEDLINE=21257802; PubMed=11358497;

RX Saito M., Takenouchi Y., Kunisaki N., Kimura S.;

"Complete primary structure of rainbow trout type I collagen

consisting of a1(I)a2(I)a3(I) heterotrimers.";

Eur. J. Biochem. 268:2817-2827(2001).

DR EMBL; AB052835; BAB55661.1; -

DR InterPro: IPR000087; Collagen.

DR InterPro: IPR000885; Fib_collagen_C.

DR Pfam: PF01410; COLFI.1.

DR Pfam: PF01391; Collagen.18.

DR ProDom: PD000093; vwc.1.

DR ProDom: PD000007; Collagen.2.

DR ProDom: PD002078; Fib_collagen_C.1.

DR PROSITE; PS01208; vwc; UNKNOWN_1.

SQ SEQUENCE 1449 AA; 137117 MW; 62EEF8A7BED652B8 CRC64;

Query Match 59.4%; Score 38; DB 13; Length 1449;

Best Local Similarity 87.5%; Pred. No. 2e+02;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 POGIAGOR 9

Db 936 POGIGCOR 943

RESULT 26

ID 09A180 PRELIMINARY; PRT: 1647 AA.

AC 09A180; (TREMBlrel. 17, Created)

DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE putative cell envelope proteinase.

OS Streptococcus pyogenes.

CC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;

CC Streptococcaceae; Streptococcus.

NCBI_TaxID=1314;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-SF370 / ATCC 700294 / SEROTYPE M1;

RX MEDLINE=21192684; PubMed=11296296;

RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.;

Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,

Qian Y., Jia H.G., Najat F.Z., Ren Q., Zhu H., Song L., White J.,

Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;

"Complete genome sequence of an M1 strain of Streptococcus pyogenes.";

Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).

DR EMBL; AE006503; AAK33444.1; -

HSP; P00782; 2SRT.

DR InterPro: IPR001899; Gram_pos_anchor.
 DR InterPro: IPR003137; PA.
 DR InterPro: IPR00209; Peptidase_S8.
 DR Pfam: PF02225; PA; 1.
 DR Pfam: PF00082; Peptidase_S8; 3.
 DR PRINTS: PR00723; SUBTILISIN.
 DR TIGRfams: TIGR01167; LpxG_anchor; 1.
 DR TIGRfams: TIGR01168; YsirK_signal; 1.
 DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
 DR PROSITE: PS00136; SUBTILASE_ASP; UNKNOWN_1.
 DR PROSITE: PS00137; SUBTILASE_HIS; UNKNOWN_1.
 DR PROSITE: PS00138; SUBTILASE_SER; 1.
 KW Envelope protein; Complete proteome.
 SQ SEQUENCE 1647 AA; 181287 MW; F36ECB965C291A2 CRC64;

Query Match
 Best Local Similarity 59.4%; Score 38; DB 16; Length 1647;
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 POGIAGORNFN 12
 DB 950 SPESFGKDNFN 961

RESULT 27
 Q909P3 PRELIMINARY; PRT; 441 AA.
 AC Q909P3; 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Ecdysone dependent glycoprotein (Fragment).
 GN SAP-R OR P110 OR BCDNA:GH08312 OR CG12070.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Galesky S., Rickoll W.L.;
 RT "Drosophila ecdysone dependent cell adhesion molecule, P110."
 RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF16129; AAD4802.1; -
 DR Flybase: FBgn0000416; Sap-R.
 DR InterPro: IPR000004; SapB.
 DR ProDom: PD001732; SapB; 3.
 DR SMART: SM00118; SAPB; 3.
 FT NON_TER 1
 SO SEQUENCE 441 AA; 49283 MW; 9372D221B33861C1 CRC64;

Query Match
 Best Local Similarity 58.6%; Score 37.5; DB 5; Length 441;
 Matches 8; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

OY 2 POGIAGO-RNF 11
 DB 247 PSGVAGCGRNF 257

RESULT 28
 Q9Y125 PRELIMINARY; PRT; 953 AA.
 AC Q9Y125; 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE SAP-R protein.
 GN SAP-R OR BCDNA:GH08312 OR CG12070.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., McKlos G.L.G.,
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S.,
 RA Borokov D., Botchan M.R., Bouck J., Brownstein P., Brotler P.,
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Crawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferrer C., Ferlito S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclob J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zheng G., Zhou Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang C.,
 RA Aghayani A., Arcalona T.T., Baxter E., Blazey R.G., Butenoff C.,
 RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
 RA Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,
 RA Houston K.A., Hummasti S.R., Kim E., Li P., Moshrefi M., Paclob J.M.,
 RA Park S., Segretra A., Sethi H., Solt E., Svirskas R.R., Weinburg T.,
 RA Ceiniker S.E.;
 RT "Full length Drosophila melanogaster cDNA sequence."
 RL Submitted (APR-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AE003775; AARS7097.1; -
 DR EMBL: AF143647; AAD38622.1; -
 DR Flybase: FBgn0000416; Sap-R.
 DR InterPro: IPR003119; SAPA.
 DR InterPro: IPR000004; SapB.
 DR Pfam: PF02199; SAPA; 1.
 DR ProDom: PD001732; SapB; 1.
 DR SMART: SM00162; SAPA; 6.
 DR SMART: SM00118; SAPB; 7.
 DR SMART: SM00118; SAPB; 7.
 SO SEQUENCE 953 AA; 105962 MW; D6CFD3E9D1502A8 CRC64;

Query Match
 Best Local Similarity 58.6%; Score 37.5; DB 5; Length 953;
 Matches 8; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

OY 2 POGIAGO-RNF 11
 DB 472 PSGVAGCGRNF 482

RESULT 29
 097550 PRELIMINARY; PRT; 84 AA.
 AC 097550;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical protein SP0563.
 GN SP0563.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 OC Streptococcaceae; Streptococcus.
 NCBI_TaxID=1313;
 RX MEDLINE=21357209; PubMed=11463916;
 RA Tettein H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
 Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.U.,
 Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
 Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
 Holtzapple E., Khouri H., Wolf A.M., Uterback T.R., Hansen C.L.,
 McDonald L.A., Feldblyum T.V., Angiuoli S., Dickson T., Hickey E.K.,
 Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
 Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
 RT "Complete genome sequence of a virulent isolate of Streptococcus
 pneumoniae."
 RL Science 293:498-506 (2001).
 DR EMBL: AE007366; AAK74719.1; -
 DR TIGR: SP0563; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 84 AA; 10142 MW; 5F3784A94EBF1BCD CRC64;
 Query Match 57.8%; Score 37; DB 16; Length 84;
 Best Local Similarity 50.0%; Pred. No. 15;
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 2 POGIAGORNF 11
 I: I: I: I: I:
 Db 75 PEGIFGERNY 84
 RESULT 30
 095061 PRELIMINARY; PRT; 328 AA.
 ID 095061;
 AC 095061;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Wasp interactor protein (Fragment).
 GN WSP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wilson L.A., Fields D., Cruz L., Friesen J., Siminovitch K.A.;
 RT "Wasp interactor protein (WISP): a novel SH3 protein that interacts
 with the Wiskott Aldrich Syndrome Protein (WASP)."
 RT Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 DR EMBL: AF001629; AAD00898.1; -
 DR InterPro: IPR001683; PX.
 DR InterPro: IPR001452; SH3.
 DR Pfam: PF00787; PX; 1.
 DR Pfam: PF00018; SH3; 1.
 DR PRINTS: PRO0452; SH3DOMAIN.
 DR PRODOM: PD000066; SH3; 1.
 DR SMART: SM00326; SH3; 1.
 DR PROSITE: PS00002; SH3; 1.
 KW SH3 domain.

FT NON_TER 328 328
 SQ SEQUENCE 328 AA; 35815 MW; 6195EF917E7485C7 CRC64;
 Query Match 57.8%; Score 37; DB 4; Length 328;
 Best Local Similarity 63.6%; Pred. No. 64;
 Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 2 POGIAGORNF 12
 I: I: I: I: I:
 Db 136 POGAGORNTN 146
 RESULT 31
 08704 PRELIMINARY; PRT; 463 AA.
 ID 08704;
 AC 08704;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE LD03583p.
 GN CG4023.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,
 RA Champagne M., Chavez C., Dorsett V., Farfan D., Fise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuno J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celinker S.;
 RT Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY069319; AAL39464.1; -
 SQ SEQUENCE 463 AA; 53154 MW; 30F6F291542247E CRC64;
 Query Match 57.8%; Score 37; DB 5; Length 463;
 Best Local Similarity 54.5%; Pred. No. 92;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SPOGIAGORNF 11
 I: I: I: I: I:
 Db 435 NPDGIAGRSNF 445
 RESULT 32
 042434 PRELIMINARY; PRT; 496 AA.
 ID 042434;
 AC 042434;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Cytokeratin 8 (Fragment).
 GN K8.
 OS Notophthalmus viridescens (Eastern newt) (Triturus viridescens).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Caudata; Salamandridae; Salamandridae;
 OC Notophthalmus.
 NCBI_TaxID=8316;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=REGENERATING LIMB BLASTEMA;
 RX MEDLINE=98075904; PubMed=9415422;
 RA Corcoran J.P., Ferretti P.;
 RT "Keratin 8 and 18 expression in mesenchymal progenitor cells of
 differentiating limbs is associated with cell proliferation and
 differentiation."
 RT Dev. Dyn. 210:355-370 (1997).
 RL -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC EMBL: AJ001296; CAA04656.1; -
 DR InterPro: IPR001664; IF.

DR InterPro: IPR003054; Keratin_II.
 DR Pfam: PF00038; filament; 1.
 DR PRINTS: PRO1276; TYPEKERATIN.
 DR PROSITE: PS00226; IF: 1.
 KW Colled coil; intermediate filament; keratin.
 FT NON_TER 1
 SQ SEQUENCE 496 AA; 5512 MW; D327911FBB7A1CBA CRC64;

Query Match
 Best Local Similarity 57.8%; Score 37; DB 13; Length 496;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 SPOGIAGORNF 12
 DB 2 SARGAPGORNFS 13

RESULT 33
 OXDP3 PRELIMINARY; PRT: 739 AA.

AC 08XDP3;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Biotin sulfoxide reductase.
 GN BISC OR Z4976 OR ECS436.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RA MEDLINE=21074935; PubMed=11206551;
 RA Pereira N.T., Plunkett G., Iiri, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Grobeck E.J., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
 RL Nature 409:529-533(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RA MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
 RA Kubura S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12";
 RL DNA Res. 8:11-22(2001).
 DR EMBL: AE005582; BAG58700.1; -;
 DR EMBL: AP002565; BAB37859.1; -;
 DR InterPro: IPR001467; Prok_Mboxred.
 DR Pfam: PF00384; Molybdopterin; 1.
 DR Pfam: PF01568; Molybdopterin; 1.
 DR TIGRFAMS: TIGR00509; bisc_fam; 1.
 DR PROSITE: PS00490; MOLYBDOPTERIN_PROK_2; 1.
 DR PROSITE: PS00932; MOLYBDOPTERIN_PROK_3; 1.
 KW Complete proteome.
 SQ SEQUENCE 739 AA; 81961 MW; 23675207DD4FE6C3 CRC64;

Query Match
 Best Local Similarity 57.8%; Score 37; DB 16; Length 739;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 SPOGIAGORNF 11
 DB 34 NPOGIRGODEF 44

RESULT 34
 O8U724 PRELIMINARY; PRT: 771 AA.

AC 08U724;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE ATP-dependent DNA ligase.
 GN ATU4632 OR AGR_L_502
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitaajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boye D., Sr.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet M., Grant C.,
 RA Kutayavon T., Levy R., Li M.-J., McClelland E., Palmeri A.,
 RA Raymond C., Rouse S., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Neeter E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
 C58";
 RL Science 294:2317-2323(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattling S., Miller N., Blanchard M.,
 RA Oucollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
 RA Houmel K., Gordon J., Vaudin M., Jarclouk O., Bpp A., Liu F.,
 RA Wollman C., Allinger M., Doughly D., Scott C., Lappas C., Markelz B.,
 RA Llanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
 RA Cleto C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 Agrobacterium tumefaciens C58";
 RL Science 294:2323-2328(2001).
 DR EMBL: AE009391; AAL45426.1; -;
 DR EMBL: AE008224; AAK88815.1; -;
 KW Ligase; Complete proteome.
 SQ SEQUENCE 771 AA; 85604 MW; BE8B7F00DA84207B CRC64;

Query Match
 Best Local Similarity 70.0%; Score 37; DB 16; Length 771;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 POGIAGORNF 11
 DB 594 PEGIEGOREF 603

RESULT 35
 O9V6K7 PRELIMINARY; PRT: 865 AA.

AC 09V6K7;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE CG4023 protein.
 GN CG4023.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY.
 RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H.C., Blake J.G., Champ E.M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abrell J.F., Abdayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokov D., Botchan M.R., Brokstein P., Brotter P.,
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strung R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RA "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003819; AAF58416.1; -;
 DR FLYBase: FBgn003801; CG4023.
 DR InterPro: IPR000461; Alpha_amyase.
 DR InterPro: IPR004193; Isoamylase.
 DR Pfam: PF00128; alpha-amyase; 1.
 DR Pfam: PF02922; isoamylase; 1.
 SQ SEQUENCE 865 AA; 99760 MW; 9B24CB9B9F01518A CRC64;

Query Match 57.8%; Score 37; DB 5; Length 865;
 Best Local Similarity 54.5%; Pred. No. 1.8e+02;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPOGIAGORNF 11
 Db 657 NPGYAGRSNF 667

RESULT 36
 ID 093EC2 PRELIMINARY; PRT; 1158 AA.
 AC 093EC2;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Impl.
 GN IMPL.
 OS Rhizobium leguminosarum (biovar trifolii).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_Taxid=386;
 OX [1]
 RP SEQUENCE FROM N.A.
 RA Bladegroen M.R., Badelt K., Stronk O.P., Lugtenberg E.J.J.,
 RA Spaik H.P.,
 RT "An avirulence locus of a symbiotic Rhizobium leguminosarum strain is

RT involved in temperature-dependent protein secretion.";
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF361470; AAL17784.1; -;
 SQ SEQUENCE 1158 AA; 127575 MW; 3374E4FBA62B5C5 CRC64;

Query Match 57.8%; Score 37; DB 2; Length 1158;
 Best Local Similarity 66.7%; Pred. No. 2.4e+02;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SPOGIAGORNF 12
 Db 170 SVQIGGTRNCN 181

RESULT 37
 ID 091717 PRELIMINARY; PRT; 1486 AA.
 AC 091717;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Alpha-1 type II collagen.
 GN COL2A1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_Taxid=8355;
 OX [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE:92011898; PubMed:1918153;
 RA Su M.W., Suzuki H.R., Bleker J.J., Solursh M., Ramirez F.,
 RT "Expression of two nonallelic type II procollagen genes during *Xenopus*
 RT laevis embryogenesis is characterized by stage-specific production of
 RT alternatively spliced transcripts.";
 RL J. Cell Biol. 115:565-575(1991).
 DR EMBL: M63595; AAA49678.1; -;
 DR InterPro: IPR000087; Collagen.
 DR InterPro: IPR000885; Fib_collagen_C.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF01410; COLF1; 1.
 DR Pfam: PF01391; Collagen; 18.
 DR Pfam: PF00093; VWC; 1.
 DR ProDom: PD000007; Collagen; 4.
 DR ProDom: PD002078; Fib_collagen_C; 1.
 DR SMART: SM00038; COLF1; 1.
 DR SMART: SM00214; VWC; 1.
 DR PROSITE: PS01208; VWF; 1.
 KW Collagen.
 SQ SEQUENCE 1486 AA; 142263 MW; 4AA95772341042F CRC64;

Query Match 57.8%; Score 37; DB 13; Length 1486;
 Best Local Similarity 75.0%; Pred. No. 3.2e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 POGIAGOR 9
 Db 975 POGISGOR 982

RESULT 38
 ID 045769 PRELIMINARY; PRT; 369 AA.
 AC 045769;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE T10C6.7 protein.
 GN T10C6.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_Taxid=6239;


```

RN [1]
RX SEQUENCE FROM N.A.
RA McMurray A.A.;
RN Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RP [12]
RX SEQUENCE FROM N.A.
RA MEDLINE-99069613; PubMed-9851916;
RX none;
RT "Genome sequence of the nematode C.elegans: A platform for
  investigating biology.";
RL Science 283:2012-2018(1998).
DR EMBL:2933388; CAB07663.1; -
DR InterPro: IPR001810; F-box; 1.
DR Pfam: PF00646; F-box; 1.
DR PROSITE: PS50181; FBOX; 1.
SQ SEQUENCE 369 AA; 43506 MW; 00DEC66EAF291041 CRC64;

Query Match
Best Local Similarity 57.0%; Score 36.5; DB 5; Length 369;
Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 SPOGIAGORNFN 12
Db 331 SPOGIAGORN-FN 341

RESULT 39
ID 09Y3P3 PRELIMINARY; PRT; 100 AA.
AC 09Y3P3;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE BK984G1.1 (Supported by EGENES) (Fragment).
GN BK984G1.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Matchews L.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL031186; CAB41538.1; -
DR InterPro: IPR000087; Collagen.
DR Pfam: PF01391; Collagen; 1.
FT NON_TER 1
SQ SEQUENCE 100 AA; 10473 MW; 220ED333CCA05668 CRC64;

Query Match
Best Local Similarity 56.2%; Score 36; DB 4; Length 100;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 POGIACOR 9
Db 8 POGIACOR 15

RESULT 40
ID 0915P4 PRELIMINARY; PRT; 149 AA.
AC 0915P4;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Probable type II secretion system protein.
GN PA0681.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxId=287;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-ATCC 15692 / PA01;
RX MEDLINE-20437337; PubMed-10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brickman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,
RA Gardner R.L., Goltzy L., Tolentino E., Westbrook-Wedman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
  opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -1- FUNCTION: INVOLVED IN A GENERAL SECRETION PATHWAY (GSP) FOR THE
CC EXPORT OF PROTEINS (BY SIMILARITY).
CC -1- SUBUNIT: THE PILI ARE POLAR FLEXIBLE FILAMENTS OF ABOUT 5.4
CC NANOMETERS DIAMETER AND 2.5 MICROMETERS AVERAGE LENGTH; THEY
CC CONSIST OF ONLY A SINGLE POLYPEPTIDE CHAIN ARRANGED IN A HELICAL
CC CONFIGURATION OF FIVE SUBUNITS PER TURN IN THE ASSEMBLED PILUS (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE N-ME-PHE PILI FAMILY.
CC -1- SIMILARITY: BELONGS TO THE PILG/OUTG/XPSG/EXES/XCPT FAMILY.
DR EMBL: AE004503; AAC04070.1; -
DR InterPro: IPR000983; Bac_GSPG.
DR InterPro: IPR002416; Bac_GSPH.
DR InterPro: IPR001120; Prok_N_methyltn.
DR PRINTS: PR00813; BACTERIALGSPG.
DR PRINTS: PR00885; BACTERIALGSPH.
DR PROSITE: PS00409; PROKAR_NTER_METHYL. 1
DR Fimbrin; Methylation; Transport; Complete proteome.
SQ SEQUENCE 149 AA; 16219 MW; 0D3E6315BC63BC28 CRC64;

```

```

QY 1 SPOGIAGORNF 11
Db 9 SPOGIAGORNF 19

Query Match
Best Local Similarity 56.2%; Score 36; DB 16; Length 149;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 SPOGIAGORNF 11
Db 9 SPOGIAGORNF 19

RESULT 41
ID 0920V4 PRELIMINARY; PRT; 409 AA.
AC 0920V4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE UDP-Gal:betaGalNAc beta 1,3-galactosyltransferase II (Fragment).
GN B3GT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-CAST/EI;
RA Liu Y., Kitano T., Kolde T., Shirolshi T., Moriwaki K., Saitou N.;
RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
  Five Mus musculus subspecies.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB039147; BAB88671.1; -
DR InterPro: IPR002659; Galactosyl_T.
DR Pfam: PF01762; Galactosyl_T; 1.
DR Glycosyltransferase; Transferase.
FT NON_TER 1
SQ SEQUENCE 409 AA; 47247 MW; 2423654F23A8F0E3 CRC64;

Query Match
Best Local Similarity 56.2%; Score 36; DB 11; Length 409;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPOGIAGORN 10
Db 1 SPOGIAGORN 10

```

Db 99 SPOGVTGLON 108

RESULT 42

Q920V3 PRELIMINARY; PRT; 409 AA.
 AC Q920V3;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase II (Fragment).
 GN B3GT2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HMI/MSF;
 RA Liu Y., Kitano T., Koide T., Shirosaki T., Moriwaki K., Saitou N.;
 RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
 RT Five Mus musculus subspecies.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB039151; BAB68675.1; -;
 DR InterPro: IPR002659; Galactosyl_T.
 DR Pfam: PF01762; Galactosyl_T.1.
 KW Glycosyltransferase; Transferase.
 FT NON_TER 1
 FT NON_TER 409
 SQ SEQUENCE 409 AA; 47220 MM; 2D2BEC4D68E8463D CRC64;

Query Match

Best Local Similarity 56.2%; Score 36; DB 11; Length 409;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPOGIAGORN 10
 ||||: 1:1
 Db 99 SPOGVTGLON 108

RESULT 43

Q920V2 PRELIMINARY; PRT; 409 AA.
 AC Q920V2;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase II (Fragment).
 GN B3GT2.
 OS Mus spicilegus (Steppe mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ZBN;
 RA Liu Y., Kitano T., Koide T., Shirosaki T., Moriwaki K., Saitou N.;
 RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
 RT Five Mus musculus subspecies.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB039153; BAB68677.1; -;
 DR InterPro: IPR002659; Galactosyl_T.
 DR Pfam: PF01762; Galactosyl_T.1.
 KW Glycosyltransferase; Transferase.
 FT NON_TER 1
 FT NON_TER 409
 SQ SEQUENCE 409 AA; 47383 MM; 0DC51AB6335A58DE CRC64;

Query Match 56.2%; Score 36; DB 11; Length 409;
 Best Local Similarity 60.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPOGIAGORN 10

Db 99 SPOGVTGLON 108

RESULT 44

Q91V58 PRELIMINARY; PRT; 409 AA.
 AC Q91V58;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase II (Fragment).
 GN B3GT2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BLG2/MSF, AND NUL/MSF;
 RA Liu Y., Kitano T., Koide T., Shirosaki T., Moriwaki K., Saitou N.;
 RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
 RT Five Mus musculus subspecies.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB039146; BAB68670.1; -;
 DR EMBL: AB039152; BAB68676.1; -;
 DR InterPro: IPR002659; Galactosyl_T.
 DR Pfam: PF01762; Galactosyl_T.1.
 KW Glycosyltransferase; Transferase.
 FT NON_TER 1
 FT NON_TER 409
 SQ SEQUENCE 409 AA; 47196 MM; 06D262B08B90F4 CRC64;

Query Match 56.2%; Score 36; DB 11; Length 409;
 Best Local Similarity 60.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SPOGIAGORN 10
 ||||: 1:1
 Db 99 SPOGVTGLON 108

RESULT 45

Q91V58 PRELIMINARY; PRT; 409 AA.
 AC Q91V58;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase II (Fragment).
 GN B3GT2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/10SNJ, BFW/2MSF, AND PGN2;
 RA Liu Y., Kitano T., Koide T., Shirosaki T., Moriwaki K., Saitou N.;
 RT "Conspicuous Differences among Gene Genealogies of 21 Nuclear Genes of
 RT Five Mus musculus subspecies.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB039144; BAB68668.1; -;
 DR EMBL: AB039145; BAB68669.1; -;
 DR EMBL: AB039149; BAB68673.1; -;
 DR InterPro: IPR002659; Galactosyl_T.
 DR Pfam: PF01762; Galactosyl_T.1.
 KW Glycosyltransferase; Transferase.
 FT NON_TER 1
 FT NON_TER 409
 SQ SEQUENCE 409 AA; 47233 MM; 9DBE67873273F13A CRC64;

Query Match 56.2%; Score 36; DB 11; Length 409;
 Best Local Similarity 60.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SPOGACORN 10
1111:1:1

Db 99 SPOGVTGLON 108

RESULT 46

Qy 091V19 PRELIMINARY; PRT; 409 AA.

AC 091V19

DT 01-DEC-2001 (Tremblrel. 19, Created)

DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)

DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)

DE UDP-Gal:betaglcNac beta 1,3-galactosyltransferase II (Fragment).

GN B3GT2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MSM/MSF; AND SWN/MSF;

RA Liu Y., Kitanu T., Koide T., Shiroishi T., Moriaki K., Saitou N.;

RT "Conspicuous differences among Gene Genealogies of 21 Nuclear Genes of
Five Mus musculus subspecies."

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB039148; BAB68672.1; -

DR EMBL; AB039150; BAB68674.1; -

DR InterPro: IPR002659; Galactosyl_T.

DR Pfam: PF01762; Galactosyl_T; 1.

KW Glycosyltransferase; transferase.

FT NON_TER 1

FT NON_TER 409

SEQUENCE 409 AA; 47206 MW; D/BD514EE40A0157 CRC64;

Query Match
Best Local Similarity 60.0%; Score 36; DB 11; Length 409;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SPOGACORN 10
1111:1:1

Db 99 SPOGVTGLON 108

RESULT 47

Qy 054905 PRELIMINARY; PRT; 422 AA.

AC 054905

DT 01-JUN-1998 (Tremblrel. 06, Created)

DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)

DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)

DE UDP-galactose:beta-N-acetyl-glucosamine-beta-1,3-galactosyltransferase
II (EC 2.4.1.-) (UDP-GAL:BETAGLCNAC BETA 1,3-GALACTOSYLTRANFERASE-II)
(BETA3GALT-II).

GN B3GT2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.

RC STRAIN=129/SVJ;

RA Henner T., Dinter A., Kuhnert P., Mattu T.S., Rudd P.M., Berger E.G.;

RT "Genomic cloning and expression of three murine UDP-galactose: beta-N-
acetylglucosamine beta1,3-galactosyltransferase genes."

RL J. Biol. Chem. 273:58-65(1998).

-1- FUNCTION: THIS PROTEIN IS RESPONSIBLE FOR THE SYNTHESIS OF
COMPLEX-TYPE N-LINKED OLIGOSACCHARIDES IN MANY GLYCOPROTEINS A AS
WELL AS THE CARBOHYDRATE MOIETIES OF GLYCOLIPIDS. GALT-I, GALT-II
AND GALT-III ENZYMES ACT ON THE SAME ACCEPTOR SUBSTRATES BUT WITH

CC DIFFERENT AFFINITY.

CC -1- CATALYTIC ACTIVITY: UDP-GALACTOSE + N-ACETYL-BETA-D-GLUCOSAMINYL-
GLYCOPETIDE -> UDP + BETA-D-GALACTOSYL-1,3-N-ACETYL-BETA-D-
GLUCOSAMINYLGLYCOPETIDE.

CC -1- COFACTOR: MANGANESE.

CC -1- PATHWAY: GLYCOSYLATION.

CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI.

CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN AND HEART. BARELY DETECTED
IN OVARY, COLON AND LYMPH NODES. NO EXPRESSED IN TESTIS, UTERUS,
CC STOMACH, KIDNEY, LIVER, LUNG, BONE MARROW, THYMUS AND SPLEEN.

CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-14 IS THE INITIATOR.

DR EMBL; MG11349461; B39612.

DR InterPro: IPR002659; Galactosyl_T.

DR Pfam: PF01762; Galactosyl_T; 1.

KW Transferase; Glycosyltransferase; Transmembrane; Signal-anchor;

KW Glycoprotein; Golgi stack; Manganese

FT DOMAIN 1 24

FT TRANSMEM 25 43

FT DOMAIN 44 422

FT CARBOHYD 75 75

FT CARBOHYD 98 98

FT CARBOHYD 119 119

FT CARBOHYD 176 176

FT CARBOHYD 226 226

FT N-LINKED (GLCNAC. . .) (POTENTIAL).

FT N-LINKED (GLCNAC. . .) (POTENTIAL).

SEQUENCE 422 AA; 49107 MW; D03D18A7484DCB03 CRC64;

Query Match
Best Local Similarity 60.0%; Score 36; DB 11; Length 422;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SPOGACORN 10
1111:1:1

Db 105 SPOGVTGLON 114

RESULT 48

Qy 09P8N1 PRELIMINARY; PRT; 436 AA.

AC 09P8N1

DT 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)

DE Cellulohydrolase.

OS Coriolus versicolor.

OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;

OC Aphyllophorales; Coriolus.

OX NCBI_TaxID=57466;

RN [1]

RP SEQUENCE FROM N.A.

RA Novo C., Clemente A., Simoes F., Mendonca D., Matos J.;

RT "Coriolus versicolor cellulohydrolase-like cDNA sequence."

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF233583; AAF35251.1; -

DR HSSP; P00725; ZCBH.

DR InterPro: IPR000254; CBD_Fungal.

DR InterPro: IPR001524; GH_6.

DR Pfam: PF00734; CBM_1; 1.

DR Pfam: PF01341; Glyco_hydro_6; 1.

DR PRINTS; PRO0733; GLHIDRLASE6.

DR ProDom; PD001821; CBD_Fungal; 1.

DR ProDom; PD00236; fcbd; 1.

DR SMART; SM00236; fcbd; 1.

DR PROSITE; PS00562; CBD_FUNGAL; 1.

DR PROSITE; PS00655; GLYCOSYL_HYDROL_F6_1; 1.

DR PROSITE; PS00656; GLYCOSYL_HYDROL_F6_2; 1.

KW Hydrolase.

SEQUENCE 436 AA; 46170 MW; E0ECC5AFBCA324D4 CRC64;

Query Match
Best Local Similarity 77.8%; Score 36; DB 3; Length 436;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 OGIAGORNF 11
 11111111
 DB 342 OGRAGQGNF 350

RESULT 49

O96A84 PRELIMINARY; PRT; 441 AA.
 ID O96A84
 AC O96A84;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Putative emul protein precursor.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN
 RP SEQUENCE FROM N.A.
 RA Leimister C., Gessler M.;
 RT "Expression of Emu genes during mouse embryogenesis.";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ416090; CAC94777.1;
 DR InterPro: IPR000087; Collagen.
 DR Pfam: PF01391; Collagen; 2.
 KM Signal.
 FT SIGNAL
 SQ SEQUENCE 441 AA; 45292 MW; 676D491C4083E18E CRC64;

Query Match 56.2%; Score 36; DB 4; Length 441;
 Best Local Similarity 87.5%; Pred. No. 1.4e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 POGIAGOR 9
 11111111
 DB 349 POGSAGOR 356

RESULT 50
 O9KMS5 PRELIMINARY; PRT; 491 AA;
 ID O9KMS5
 AC O9KMS5;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Amc.
 GN AMNC.
 OS Pseudomonas sp.
 OC Bacteria; Proteobacteria.
 OX NCBI_TaxID=306;
 RN
 RP SEQUENCE FROM N.A.
 RA STRAIN-AP-3;
 RA Takenaka S., Murakami S., Kim Y.J., Aoki K.;
 RT "Complete nucleotide sequence and functional analysis of the genes for
 RT 2-aminophenol metabolism from Pseudomonas sp. AP-3.";
 RL Arch. Microbiol. 174:265-272(2001).
 RN
 RP SEQUENCE FROM N.A.
 RA STRAIN-AP-3;
 RA Takenaka S., Murakami S., Shinke R., Hatakeyama K., Yukawa H.,
 RT "Novel genes encoding 2-aminophenol 1,6-dioxygenase from Pseudomonas
 RT sp. AP-3 growing on 2-aminophenol and catalytic properties of the
 RT purified enzyme.";
 RL J. Biol. Chem. 272:14727-14732(1998).
 CC -1- SIMILARTY: BELONGS TO THE ALDEHYDE DEHYDROGENASES FAMILY.
 DR EMBL; AB020521; BAB03533.1;
 DR HSSP; P56533; 1A4S.
 DR InterPro: IPR002086; Aldehyde_dehydr.
 DR Pfam: PF00171; aldehyd. 1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.
 DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.

KW Oxidoreductase.
 SQ SEQUENCE 491 AA; 53722 MW; 1B196D37FF534551 CRC64;

Query Match 56.2%; Score 36; DB 2; Length 491;
 Best Local Similarity 70.0%; Pred. No. 1.5e+02;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 POGIAGORNF 11
 11111111
 DB 103 POGIANFRNF 112

Search completed: May 16, 2003, 10:41:03
 Job time : 50 secs